Repurposing antiretroviral drugs for treating triple-negative breast cancer via LINE-1 regulation

Pey-Tsyr Chiou

August 2019

A thesis submitted for the degree of Doctor of Philosophy
of the Australian National University

The John Curtin School of Medical Research
ANU College of Health & Medicine
The Australian National University
Canberra, Australia

Statement of originality

The contents presented in this thesis are my original work and have not been written or published by another person except where otherwise acknowledged.

Pey-Tsyr Chiou

August 2019

Acknowledgments

I would like to express my deepest gratitude to my supervisors, Dr Danny Rangasamy and Associate Prof Marco Casarotto, for the valuable suggestions and support that they gave me for this project as well as my PhD life at the ANU. I am grateful to Danny for introducing me to the research field of cancer therapeutics and for giving me the freedom to direct my project. With his help, I have had a chance to gain insight into the academic world. I cannot express my gratitude enough to Marco. He is one of the best supervisors I have worked with and it was a pleasure to be supervised by him.

I would also like to thank all the people and facilities that have assisted me during this project, including the ACRF Biomolecular Resource Facility, the Imaging & Cytometry Facility, and many PhD students at JCSMR (the list is too long to mention everyone here, but I will always remember them). They gave me lots of support in my research and thesis writing. I would like to express special thanks to Dr Stephen Ohms - he is extraordinarily passionate about his work and was a great mentor, showing me how you should be as a scientist. This thesis was copyedited by Professional Editor Kathie Brown of A+ Academic Editing Canberra and I thank for her help as well.

I am also greatly indebted to all the members in my previous and current laboratories: Christine, Dmitry, Melanie, Sandali, Shouvik, and Srinivasan. They were very supportive and a friendly, lovely, warm group of people to work with – I wish I had met them much earlier. I will forever be grateful to have worked with them.

Finally, I would like to give my deepest thanks to my family – my mom, dad, sister, and brother, and also to my partner, Ko-Li. I could not have completed this work without their support. I will always appreciate their love and never forget what they have done for me.

Abstract

The most common cancer in women is breast cancer with approximately 1 in 8 women developing this disease in their lifetime. Clinically, breast cancer can be divided into distinct subtypes based on the presence or absence of hormone receptors such as estrogen receptor (ER), progesterone receptor (PR) and expression of the HER2 gene. In this project, one of the most difficult to treat subclasses - triple-negative ($ER^-/PR^-/HER2^-$) breast cancer (TNBC) is studied. Until now, very limited drug treatment strategies are available for TNBCs because of a lack of hormone receptors as potential drug targets. Thus, it is a matter of urgency to seek specific treatment tailored to TNBC patients. This raises the prospect that antiretroviral drugs might be able to repurpose as anticancer drugs for TNBCs. In the mid-1990s, the incidence of the AIDS-related cancers was greatly reduced in the HIV patients due to the introduction of antiretroviral therapy to the HIV patients. It has been suggested that a direct inhibitory effect of antiretroviral drugs on the reverse transcriptase activity of long interspersed nuclear element 1 (LINE-1) in tumour cells could be a crucial factor. LINE-1 is the most important transposon with autonomous retro-transposition ability in humans. It can alter gene regulations and cause somatic mutations. Since LINE-1 has the potential to adversely affect individuals, it is silenced in differentiated tissues by diverse endogenous mechanisms. Nonetheless, LINE-1 is highly expressed in many cancers, especially in breast carcinomas. Therefore, it is worth examining whether antiretroviral drugs can be repurposed as anticancer drugs for treating TNBCs and to further understand the relationship between LINE-1 and TNBCs. Antiretroviral drug-induced anticancer effects may be relevant to down-regulation of the fatty acid metabolism pathway. In this project, two antiretroviral drugs - Efavirenz and SPV122, had been shown to cause cell death and cell proliferation retardation, thus, effectively eliminating cancer cells in a range of TNBC cell lines. Additionally, LINE-1 suppression had been observed in the antiretroviral drugs-treated breast cancer cell lines implying a potential link between LINE-1 and TNBCs. Whole-genome RNA sequencing data further highlighted the possible mechanisms involved in this anticancer process. It seemed the fatty acid metabolism pathway could be a key regulator in this anticancer process. Many key genes involved in fatty acid metabolism were down-regulated after drug treatments. However, the antiretroviral drugs-treated MCF10AT and MCF10CA1α cells were found to present some mesenchymal markers which are often characteristic signs of poor prognostic outcomes thereby highlighting the complexity of TNBCs.

The RNA sequencing data also strongly implied that cancer stem cells (CSCs) could play a role in these confusing results. CSCs are a small group of cancer cells with stem cell-like abilities, and they are thought to be responsible for drug resistance, cancer metastasis, and cancer recurrence. Interestingly, in a series of experiments, different groups of CSCs showed various responses to a range of drugs. $ALDH^{high}$ epithelial-type CSCs were significantly reduced after antiretroviral drug treatment; whereas, $CD44^{+}/CD24^{-}$ mesenchymal-type CSCs were increased after treatment. These results highlighted the importance of CSC heterogeneity and implied that mesenchymal-type CSCs have greater resistance to the drugs than other cancer cells. Finally, the functional CSC assay demonstrated that CSCs can be eliminated by the antiretroviral drugs indicating that Efavirenz and SPV122 might be able to target both non-CSCs and CSCs. To combine all the results together, Efavirenz and SPV122 could potentially be valid anticancer drugs for treating TNBCs by regulating cancer fatty acid metabolism. Follow-up experiments are necessary to further understand how antiretroviral drugs impact anticancer processes.

Publication Arising from This Study

Chiou, P., Ohms, S., Casarotto, M., and Rangasamy, D. (2017). *The role of L1 retrotransposons in triple-negative breast cancer*. Poster presented at 3rd World Congress on Controversies in Breast Cancer; 2017 October 26-28; Tokyo, Japan.

Chiou, P., Ohms, S., Casarotto, M., and Rangasamy, D. (2017). *The Therapeutic Potential of NNRTIs in Breast Cancer Stem Cells.* Poster presented at Frontiers in Cancer Science; 2017 November 6-8; Singapore.

Chiou, P., Ohms, S., Casarotto, M., and Rangasamy, D. (2017). *The role of L1 in Breast Cancer.* Poster presented at The ASMR New Investigator Forum; 2017 June 7; Canberra, Australia.

Chiou, P. and Rangasamy, D. (2015). *Anti-retroviral Efavirenz drug reduces DNA double-strand breaks in breast cancer*. Poster presented at 1st Australian Cancer and Metabolism Meeting; 2015 April 29-May 1; Sydney, Australia.

Table of Contents

Table of Figures

Chapter 1 General introduction

1.1 Introduction

Breast cancer is the most common cancer diagnosed in Australian women (AIHW, 2017) with incidences predicted to rise worldwide in the coming decades (WHO, 2014). Although its early detection and new therapeutic strategies have increased the survival rate of breast cancer patients (AIHW, 2017), approximately 10% of patients still die within five years of diagnosis (AIHW, 2017). The increase in breast cancer mortality depends on many factors including cancer metastasis, cancer recurrence, and the origin of the cancer, all of which are challenging for the design of future breast cancer therapeutics. Current therapeutic treatments have unpleasant side effects which may impact patients' daily lives. Thus, the goal for breast cancer researchers is to develop novel cancer therapies which increase both the survival rate and the quality of life for breast cancer patients.

In this introduction, I will focus on breast cancers, retrotransposons, and antiretroviral drugs. More specifically, I will describe triple-negative breast cancers (TNBCs); the role of cancer stem cells (CSCs) in cancer biology; the characteristics of different types of transposons and their function in the human genome. The utility and mechanisms of antiretroviral drugs will be detailed since they appear promising candidates as anti-breast cancer drugs. This will be followed by current evidence and hypotheses about the influence of transposons, in particular, long interspersed nucleotide element 1 (LINE-1), in epithelial cancers. This introduction should help readers to understand the relationships between LINE-1, antiretroviral drugs, and novel TNBC treatments. Furthermore, it explains how the cancer field may benefit from these findings.

1.2 Breast cancer

'Cancer is a disease that results from the successive accumulation of genetic and epigenetic alterations in oncogenes and tumour-suppressor genes, leading to uncontrolled cell growth' (Virani et al., 2012).

Breast cancer continues to be a growing public health concern worldwide. According to a recent Australian government report, breast cancer transcends colon cancer and became the most common cancer in females (AIHW, 2017). The last few decades have seen a growing trend in the increased incidence of breast cancer, possibly due in part to an increase in life expectancy and an improvement in breast cancer detection techniques (WHO, 2014). On average, around 50 people are diagnosed with breast cancer every day in Australia (AIHW, 2017). It is the fourth leading cause of cancer-related deaths in females in Australia (AIHW, 2017). Thus, developing more effective treatment strategies for breast cancer patients is an urgent priority.

Table 1-1 Staging of breast cancer

Stage	Description
Stage 0	Carcinoma *in situ*
Stage I	Early-stage, non-invasive, no metastasis, size < 1cm
Stage II	Early-stage, non-invasive or slightly invasive, no metastasis, larger size
Stage III	Advanced breast cancer, invasive, no metastasis, larger size or any size with ulceration or skin nodules
Stage IV	Advanced breast cancer, invasive, metastasis, any size

Breast Cancer Treatment. (2019, January 31). Retrieved February 8, 2019, from https://www.cancer.gov/types/breast/hp/breast-treatment-pdq#link/_695_toc)

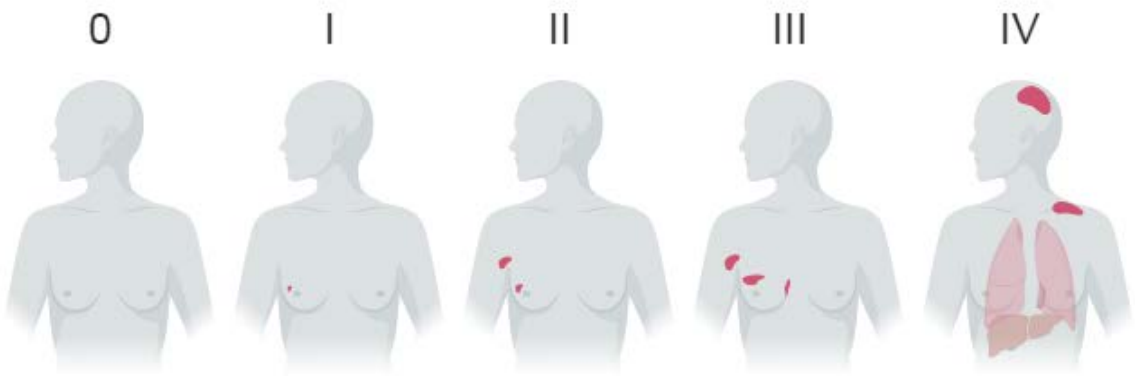


Figure 1-1 Different stages of breast cancer
The numbers represent breast cancer stages; the red shadows represent spreading areas of breast cancer in different stages.

Breast cancer treatment strategies depend on the type of breast cancer and the stage of the disease upon detection. There are five major stages for describing the progression of breast cancer (Table 1-1 and Figure 1-1). Stage 0 describes non-invasive breast cancers; stage I describes slightly invasive breast cancers of smaller size; stage II describes more invasive breast cancers of larger sizes which can affect surrounding lymph nodes; stage III describes larger tumours affecting many lymph nodes; and Stage IV describes invasive and metastatic breast cancers with the worst prognostic outcomes (NIH, 2018). Breast cancers detected in the early stages are generally much easier to treat than breast cancers detected in later stages; however, the type of breast cancer is also important in determining the prognostic outcome.

There are many subclasses of breast cancers identified by their characteristics and gene expression. Despite the monoclonal origins of solid tumours, they are composed of a heterogeneous population of cells with different proliferative, differentiate, and tumour-initiating potential (Virani et al., 2012). One of the most recognised and accepted classification methods uses hormone receptor identification. From an immunohistochemical perspective, breast cancer can be separated into four major groups: the luminal A breast cancers, the luminal B breast cancers, the human epidermal growth factor receptor 2 positive ($HER2^+$) breast cancers, and the triple-negative breast cancers (TNBCs) (Onitilo et al., 2009). While the luminal and the $HER2^+$ breast cancers are usually treated by specifically targeting their hormone receptors, systemic chemotherapies or radiotherapies are more often used for eliminating TNBC tumour tissues (Denkert et al., 2017).

Targeted therapy is a fundamental and indeed a commonly used approach for treating cancers: chemotherapy and radiotherapy being the most potent anticancer strategies employed. However, quite a few side effects including vomiting, nausea, and hair loss have been reported after such treatments (Coates et al., 1983). Moreover, although the success rate of chemotherapies is high in eliminating TNBCs, some patients can relapse within three years of the primary treatment (Foulkes et al., 2010, Hudis and Gianni, 2011). One of the greatest challenges for treating breast cancers is determining how to target TNBCs since current targeted therapy methods depend upon the presence of hormone receptors. Novel therapies which specifically target TNBC tissue, while avoiding normal tissues, are needed to increase the quality of life for patients.

1.2.1 Triple-negative breast cancers

Triple-negative breast cancer refers to breast cancers lacking the estrogen receptor (ER), the progesterone receptor (PR) and *Her2/neu* expression (Collignon et al., 2016). They account for a significant percentage (15% to 25%) of breast cancer cases (Reddy, 2011). Treatment of TNBCs is more challenging than non-TNBCs because current hormone-based therapies mainly target receptors which are either absent or poorly expressed. Moreover, TNBCs are often characterised by less differentiation, higher proliferation rates, and larger tumour sizes than non-TNBCs. Therefore, more TNBCs are identified as invasive ductal carcinomas and frequently metastasise (Kalimutho et al., 2015). Furthermore, TNBCs can be particularly aggressive and are more likely to recur in the early stages than other breast cancer subtypes (Foulkes et al., 2010). Patients with TNBC have a lower five-year survival rate than the average breast cancer five-year survival rate (Foulkes et al., 2010, Hudis and Gianni, 2011). During TNBC treatment, prevention of cancer recurrence and cancer metastasis have become central issues in the maintenance of patients' cancer-free status. There is an increasing concern that some traditional therapies are ineffective when dealing with metastatic cancers (Zeichner et al., 2016). Thus, improved diagnosis and therapeutic approaches for TNBCs are urgently needed.

Currently, TNBC patients are disadvantaged due to limited treatment options. As observed by Kalimutho et al. (2015), the lack of causally proven, high-frequency oncogenic drivers for targeting is a crucial challenge for TNBC treatment. This means that very few drugs can target TNBCs sufficiently. There are some TNBC drugs in the development phase which act by inhibiting epidermal growth factor receptor (EGFR), mammalian target of rapamycin (mTOR), phosphoinositide 3-kinase (PI3K), and Janus kinase. Alternative targets are other molecules/signalling pathways which may affect TNBC development (Kalimutho et al., 2015). However, to date, none of them has passed phase III clinical trials (Denkert et al., 2017). Moreover, until very recently, Atezolizumab (Roche) and Nanoparticle albumin-bound-paclitaxelis (Celgene) co-treatment has been approved as the first immunotherapy strategy for treating advanced or metastatic PD-L1-positive TNBC by the US Food and Drug Administration (FDA)(Schmid et al., 2018).

A significant number of patients (6%) will present with *de novo* metastatic disease, and 10% to 40% of patients with localised breast cancer will relapse systemically (Zeichner et al., 2016).

Only a few treatments are available for treating metastatic TNBC, most of which use single-agent chemotherapy such as microtubule inhibitors, anthracyclines, and antimetabolite agents (Zeichner et al., 2016). The poor drug efficacy and the occurrence of frequent visceral or central nervous system metastases result in a lower overall survival rate for metastatic TNBC patients (Zeichner et al., 2016).

There are many reasons why the development of TNBC-specific targeted therapies is proving challenging to researchers. For example, the targeted pathways might be regulated by other genes and pathways and therefore the target sites cannot be fully inhibited in humans. TNBCs were originally considered to be a homogeneous group of breast cancers sharing similar genetic backgrounds and lacking hormone receptors. Recently, however, researchers have found heterogeneity exists within the group and that many different subclasses of TNBCs can be classified. Some studies in the field of cancer biology have focused only on the general characteristics of TNBCs; however, heterogeneity should be considered since cancer cells with different genetic backgrounds may respond differently to different treatments (Lehmann et al., 2011). There are now many new techniques to help us identify the different subclasses of TNBCs, the latest of which was defined by Lehmann and his colleagues (Lehmann et al., 2011). TNBCs can be separated into *basal-like 1*, *basal-like 2*, *immunomodulatory* (IM), *mesenchymal-like*, *mesenchymal-stem like*, *luminal androgen*, and other TNBCs (Lehmann et al., 2011). Every subclass uses different signalling pathways, and therefore, may respond differently to drug treatments (Chen et al., 2012b). It is now possible to improve our understanding of how different subclasses of TNBCs might respond to various drugs, to make prognostic predictions, and to find novel targets for treating TNBCs.

An alternative explanation for the difficulties experienced by researchers in developing valid anti-TNBC drugs is therapeutic resistance. Evidence suggests that therapeutic resistance may be associated with the existence of cancer stem cells (CSCs) (Kai et al., 2010). If the treatments cannot effectively eliminate CSCs in addition to other cancer cells, the tumours may reappear rapidly after treatment (Rodriguez Salas et al., 2010).

1.2.2 Cancer stem cells

Cancer stem cells (CSCs) are cancer cells with stem cell-like abilities. This type of cancer cell has captured the attention of cancer clinicians and may be one of the key factors associated with their difficulty in treating cancers. Ever since CSCs were reported in leukaemia two decades ago, they have been attracting a lot of interest; however, many details are still under debate (Simard and Engels, 2010). CSCs are also called cancer-initiating cells because when injected into immunodeficient mice, they can generate tumours (Charafe-Jauffret et al., 2009). Numerous studies suggest that the eradication of CSCs could be among the most important strategies for developing successful cancer treatments because they impact upon cancer self-renew, cancer metastasis, drug resistance, and are responsible for tumour recurrence (Liu et al., 2015, Kai et al., 2010, Boyle and Kochetkova, 2014, Charafe-Jauffret et al., 2009). Under most conditions CSCs are quiescent, but they can be triggered to induce self-renewal activities and to reproduce progenitor cells (Chen et al., 2016). Clearly, the existence of CSCs is a major problem in treating cancers. Researchers have recently shown that if anti-cancer drugs kill only non-CSCs, CSCs will rapidly induce further cancer growth. Thus, by targeting CSCs, the issues of cancer metastasis and cancer recurrence might be addressed, and patients' lives might be extended. The discovery of CSCs has provided an insight into why cancer is sometimes very difficult to treat and eliminate. Cancer treatment regimens need to be devised to overcome technical challenges.

CSCs are difficult to characterise because there does not appear to be a universal CSC marker (Moghbeli et al., 2014, Snyder et al., 2018). Typically, different markers identify different groups of CSCs with very little overlap (Liu et al., 2014a). In breast cancer, there are several frequently used CSC markers, including CD44/CD24, ALDH, CD133, CD49f *in vitro* with more markers identified in clinical samples (Moghbeli et al., 2014). However, no universal marker represents all breast CSCs.

At least three different cellular markers and one functional assay are commonly used to identify breast CSCs (BCSCs) in breast cancer cell lines (Figure 1-2A). Mesenchymal-type CSC, also known as $CD44^{+}/CD24^{-}$ CSC, is one of the most well-known: high levels of expression of surface marker CD44 and low or no expression of surface marker CD24 characterise this kind of CSC. Epithelial-type CSC or $ALDH^{high}$ CSC is also commonly recognized. It has high

aldehyde dehydrogenase (ALDH) activity which can catalyse the oxidation of aldehyde readily. By using the ALDEFLOUR assay (Stemcell technology), epithelial-type CSC can be detected using flow cytometry. The side population (SP) is another group of less promising CSCs which use many ATP-binding cassette (ABC) transporters to pump cytotoxic materials out of the cells. Some researchers believe this allows the cells to survive in extreme conditions. These cells can be identified using Hoechst staining and flow cytometry because the levels of Hoechst remaining in SP cells are lower than other cells, allowing them to be separated from the whole population. Tumorsphere formation is used to identify another group of CSCs. Some cancer cells are able to form spheres when cultured in three-dimensional conditions of low-nutrition and are considered to be cells with CSC potential. These cells have been reported to form tumours from very small tumorspheres. Unfortunately, there is no conclusive evidence to indicate which one of these different types of CSCs is the correct definition of CSCs, and these different types of CSCs only overlapped with each other in a quite small amount. Therefore, research to date has not yet identified valid drugs for targeting CSCs and only a limited number of potential anti-CSC drugs has been investigated. More studies are therefore necessary to increase our understanding of CSCs before they can be eliminated from cancer patients.

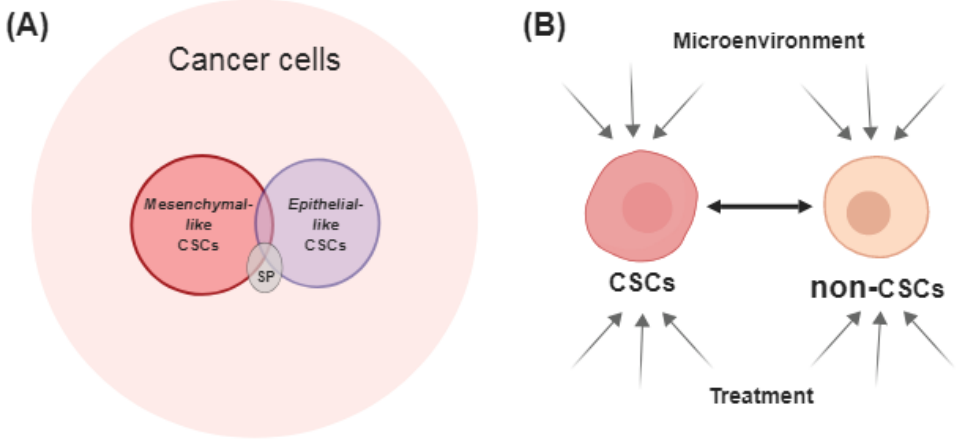


Figure 1-2 Relationships among different types of BCSCs and CSC plasticity
(A) According to various CSC identification methods, there are at least three different types of BCSCs which have very limited overlap. (B) Some non-CSC and CSC may transit their status in between non-CSC and CSC because of many different stimulations.

Another difficulty faced by researchers trying to target CSCs relates to their plasticity (Figure 1-2B). There are many studies reporting that some cancer cells can switch between non-CSCs and CSCs (Liu et al., 2014a); however, there is very little known about the factors affecting CSC plasticity. This process may be caused by environmental changes or epigenetic regulation (Chen et al., 2016). Thus, while drugs may target existing CSCs, subsequently some non-CSCs might switch to CSC status and regenerate cancers. Therefore, a drug which can eradicate both non-CSCs and CSCs together could be the next generation anti-cancer drug for improving breast cancer prognostic outcomes.

1.3 Drug repurposing

Drug repurposing is an area of drug therapeutics which is currently generating significant interest in the field of cancer biology. It is attractive since reducing the possibility of wasting time and money, and it may build up a baseline before further drug investigation. New drugs face a high failure rate, mainly because the *in vitro* conditions are very different from *in vivo* conditions (Sleire et al., 2017). However, if the drug of interest has previously been used in clinical trials for another purpose, scientists already have some baseline information for the repurposed drugs, such as possible side effects, minimal dosage, and potential drug-drug interactions which can shorten the drug discovery process (Sleire et al., 2017). One of the most successful cases of drug repurposing for treating breast cancers is Metformin (Zi et al., 2018). Metformin is a type II diabetes drug that can inhibit the mTOR pathway and can also induce cancer cell apoptosis (Zi et al., 2018).

1.3.1 Antiretroviral drugs

Retroviruses are a class of RNA virus, which use reverse transcriptase (RT) to insert their own sequence into the host DNA in order to replicate themselves using the host replication machinery. The human immunodeficiency virus (HIV) is the most well-known retrovirus, causing acquired immune deficiency syndrome or AIDS. Even though the incidence of AIDS is increasing, the mortality rate is decreasing because of the widespread use of antiretroviral drugs (Simard and Engels, 2010).

The anticancer potential of antiretroviral drugs has been the subject of research since the 1990s when scientists realised the incidence of AIDS-related cancers was dramatically reduced after treatment with antiretroviral drugs (Simard and Engels, 2010). About two decades ago, highly active antiretroviral therapy (HAART) was introduced to HIV-infected patients resulting in the reduced motility rate of AIDS-related cancers (Simard and Engels, 2010). It was considered a great success in the treatment of AIDS patients and the virus was no longer able to induce AIDS-related cancers. However, several researchers believe this success resulted from the inhibition of endogenous reverse transcription in cancers (Simard and Engels, 2010), activity which was derived from long interspersed nucleotide element 1 (LINE-1), a highly mobile transposable element within the human genome.

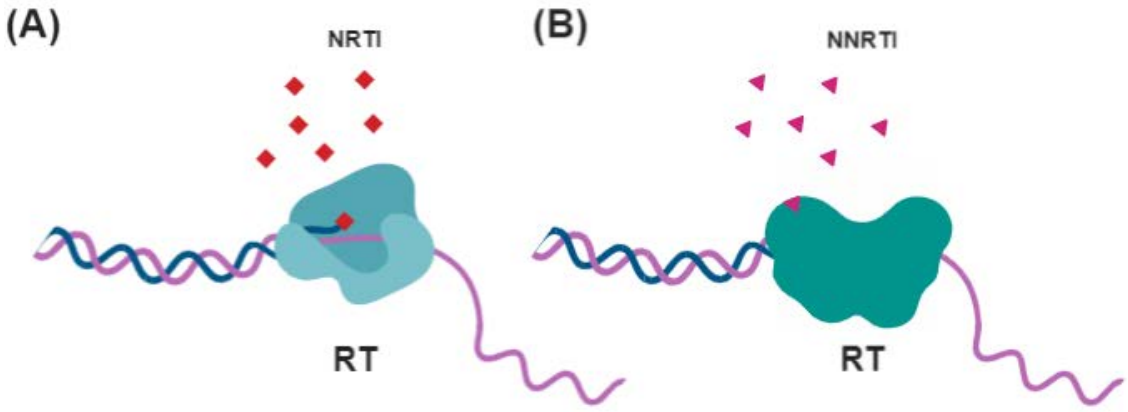


Figure 1-3 The mechanisms of RT inhibitors for treating AIDS
(A) The mechanism of NRTIs is to terminate virus DNA replication. Management of NRTI treatment. (B) The mechanism of NNRTIs is to block the function of RT and stop DNA synthesis.

RT inhibitors are the most widely used antiretroviral agents, and many of them are highly effective at stopping HIV amplification. Furthermore, they have been extensively used for treating many kinds of epithelial cancers, including acute myeloid leukaemia (AML) and prostate cancers in cell culture (Sciamanna et al., 2005). Two major groups of RT inhibitors are well-studied in HIV treatment: one group being nucleoside RT inhibitors (NRTIs), and the other being non-nucleoside RT inhibitors (NNRTIs). NRTIs are nucleoside analogues which can act as terminators of growing DNA strands (Figure 1-3A); whereas, NNRTIs can bind to a 'pocket' of RT causing conformational changes in the RT, thus inhibiting RT function (Figure 1-3B) (Bastos et al., 2016).

LINE-1 inhibition has been studied by many researchers using a retrotransposition assay (Dai et al., 2011, Peddigari et al., 2013, Heras et al., 2013). In most of the retrotransposition studies in bacterial models, some NRTIs and NNRTIs can effectively inhibit LINE-1 expression (Dai et al., 2011). There is evidence to suggest that NRTIs can more efficiently inhibit LINE-1 expression than NNRTIs in the retrotransposition assay (Jones et al., 2008); although the effects might differ depending on the cell line. Although some of these antiretroviral drugs have been confirmed to operate as LINE-1 inhibitors, their precise effects and their mechanisms of action remain unclear.

1.4 Transposable elements

Transposable elements (TEs or transposons) are a group of genes which can alter their own location in the genome and may play substantial roles in human development and evolution. TEs were previously considered as non-functioning 'junk genes' (Burns, 2017). Recently, however, there has been renewed interest in these 'junk genes', because scientists have noticed that they make up a large proportion of the human genome. The Human Genome Project identified more than 45% of the human genome as comprising TEs (Sachidanandam et al., 2001). Since the proportion of TEs is greater than the protein-coding genes (~2%), scientists thus anticipated that TEs could play a critical role in the human genome (de Koning et al., 2011). Subsequently, TEs have been a topic of great interest in a wide range of fields. Therefore, the significance of TEs in human development and evolution is a hot-button topic in the field of biology. To further understand each type of TE and their influence in humans, I describe the different types in the following sections.

1.4.1 DNA transposons

One of two major classes of TEs is the DNA transposon. In humans, DNA transposons are believed to employ a 'cut and paste' mechanism for mobilization in the genome by using their own encoded transposase. DNA transposons might therefore be cut from their original location in the chromosomal DNA and then be inserted into a new genetic location. Little is known about DNA transposons because only 2.8% of inactive DNA transposons are present in the human genome and only a small amount of their remnants have been found in the germline (Evsikov and Marin de Evsikova, 2016). However, scientists believe DNA transposons were active 37 million years ago and could be a driving force in human evolution (Cordaux and Batzer, 2009). Since DNA transposons are more like molecular fossils, appearing to play no current role in the human genome, this review will focus more on retrotransposons.

1.4.2 Retrotransposons

Retrotransposons are widespread endogenous gene fragments residing in the genome with mobility (Farkash et al., 2006). They can 'copy and paste' themselves to new genomic positions via RNA-mediated reverse transcription (Cordaux and Batzer, 2009). Retrotransposons comprise approximately 42% of the human genome, including 8.3% of long terminal repeat (LTR) elements and 33.7% of non-long terminal repeat (non-LTR) elements (Sachidanandam

et al., 2001). In contrast to DNA transposons, some active retrotransposons still have the ability to move within the genome, in particular, non-LTR elements (Cordaux and Batzer, 2009).

1.4.3 Long Terminal Repeat (LTR) elements

Endogenous retroviruses (ERVs) are the major group of LTR elements in humans and can be historically traced back to retroviral-associated infections in the genome. Their name indicates their main characteristics, and they are the remnants of ancient retroviral infections (Seifarth et al., 2005). Similar to retroviruses, they are supposed to be able to copy themselves back into new genetic locations with long terminal repeat ends; however, they have lost part of their sequence and hence their jumping function (Evsikov and Marin de Evsikova, 2016). They have also been identified as inactive TEs and exert limited influence on the human genome (Cordaux and Batzer, 2009). Some other LTR elements such as mammalian apparent LTR retrotransposons (MaLR) are considered as active 'jumping genes' in certain rodents; however, there is no evidence to suggest that MaLR elements appear in the human genome (Evsikov and Marin de Evsikova, 2016). In comparison with LTR elements, there are more copy numbers of non-LTR elements in the human genome, implying that non-LTR are more important than LTR elements.

1.4.4 Non-Long Terminal Repeat (non-LTR) elements

To date, the only TEs identified in humans, as being currently active, are non-LTR elements. Hence they are the only transposons which may have the ability to impact the human genome. Non-LTR elements, including the non-autonomous short interspersed nuclear element (SINE) and the autonomous long interspersed nuclear element (LINE), have been reported to move within the human genome (Kroutter et al., 2009). SINEs are relatively short non-autonomous sequences relying on the reverse transcription function of LINEs to assist their insertions; whereas, LINEs are much longer sequences than SINEs and can insert themselves at new gene locations autonomously (Koito and Ikeda, 2013). Embryogenesis and cell development are heavily involved in these non-LTR elements. Many non-LTR element insertions have also been confirmed as potential resources in several genetic disorder diseases (Kroutter et al., 2009).

1.4.5 Long interspersed nucleotide element 1 (LINE-1)

Long interspersed nuclear element 1 (LINE-1) is one of the most common and active autonomous retrotransposons in the human genome, comprising approximately 17% of the

genomic sequence. LINE-1 is encoded as a 40kDa RNA binding protein ORF1p and a 150kDa protein ORF2p, with the latter possessing endonuclease and reverse transcriptase activity (Figure 1-4) (Sciamanna et al., 2014). Because of its 'copy and paste' ability in the human genome, it impacts gene structure and gene expression (Koito and Ikeda, 2013). It is now well established that LINE-1 can impair the stability of the genome and is thought to be responsible for gene damage, gene insertion, and gene breakdown (Beck et al., 2010). LINE-1 reverse transcription activities can be extremely harmful to human beings, and the human body has developed many different mechanisms for preventing LINE-1 from jumping within the genome (Rodić and Burns, 2013). LINE-1 is rarely expressed in normal healthy differentiated tissues (Rodić and Burns, 2013), although LINE-1 expression has been associated with some diseases that are caused by DNA mutagenesis and instability of LINE-1 elements (Zhang et al., 2014).

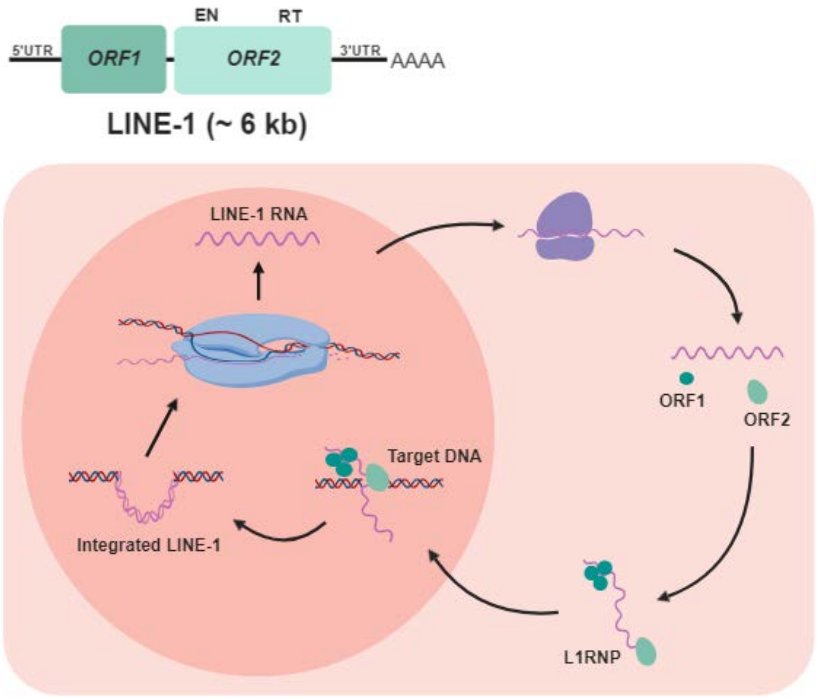


Figure 1-4 The structure and the 'copy and paste' mechanism of LINE-1
The picture on the top depicts the structure of LINE-1. It encodes an ORF1 protein and an ORF2 protein-containing endonuclease (EN) and reverse transcriptase (RT) regions. The picture on the bottom depicts the process of LINE-1 insertion.

1.4.6 LINE-1 in Cancer

The expression of the LINE-1 element is highly consistent with the occurrence of several human neoplasms, specifically lung, prostate, ovarian carcinoma as well as breast cancer (Rodić and Burns, 2013). While the mechanisms that underpin LINE-1 expression in many cancers are not fully understood, it is considered to be a potential biomarker for many epithelial cancers due to its distinct patterns of expression in cancer tissues and patients (Houédé et al., 2014, Houédé et al., 2018, Hecht et al., 2018). Normally, elements are not expressed in healthy differentiated tissues because of genome defence mechanisms such as DNA methylation and RNA interferons (Harris et al., 2010). In contrast, levels of LINE-1 are particularly high in cancer tissues, in tumour microvesicles, as well as in the plasma of breast cancer patients. LINE-1 may act to modify normal cells in the tumour microenvironment thus making them more amenable to tumour growth (Chen et al., 2012a).

Interestingly, LINE-1-encoded proteins display alternative patterns of expression in different kinds of breast cancer cell lines. The expression patterns of LINE-1 differ in non-invasive and invasive cancer cells (Chen et al., 2012a). This suggests that detecting LINE-1 expression could be useful for distinguishing the various subtypes and clinical stages of cancer cells. Moreover, survival rates of patients with a particular subclass of invasive breast cancer correlated with detection of LINE-1-encoded ORF1p and ORF2p in the nucleus but not in the cytoplasm (Chen et al., 2012a). These studies indicate that LINE-1-encoded proteins have the capability to become biomarkers for prognostic prediction in breast cancer, which may allow us to predict which patients will end up in high-risk categories. However, at present, the molecular and signalling pathway changes that occur during LINE-1 expression remain unclear. The impact of LINE-1 on epithelial cancers remains undefined in relation to their mechanism of action.

1.5 LINE-1, antiretroviral drugs, and cancers

Repurposing antiretroviral drugs as anticancer drugs using LINE-1 inhibition could be a potential therapeutic pathway for treating breast cancer patients. LINE-1 encodes an RT enzyme that may have a key role in tumorigenesis (Patnala et al., 2014). In biochemical analyses, several antiretroviral drugs have been shown to inhibit the RT activity of LINE-1. In retrotransposition assays, they have been observed to reduce the number of L1 retrotransposition events (Dai et al., 2011). Recent studies have revealed that antiretroviral drugs may effectively reduce cancer cell proliferation in prostate and pancreatic cancer (Houédé et al., 2014, Hecht et al., 2018). Inhibition of LINE-1 RT activity has also been shown to promote differentiation in several progenitors and transformed cell types (Spadafora, 2004). Some researchers have been attempting to find a link between LINE-1 RT inhibition and cancer elimination. It is now well established that LINE-1 is expressed in tumour tissues but not in normal tissues; however, the influence of LINE-1 inhibition on cancer treatment remains ill-defined.

LINE-1 inhibition is associated not only with oncogene expression (Patnala et al., 2014) but also with the regulation of certain small non-coding RNAs (Ohms and Rangasamy, 2014) which may be involved in regulating cancer progressive protein expression (Xue and He, 2014). According to a recent paper (Patnala et al., 2014), inhibiting the function of LINE-1 with Efavirenz, an antiretroviral drug, reduces the expression of some well-known oncogenes in breast cancer including EGFR, erythroblastosis leukemia viral oncogene homolog (ERBB4) and others. On the other hand, silencing LINE-1 mRNA with a specific endo-siRNA (endo453) in the T47D breast cancer cell line increases a group of tumour suppressor miRNAs. These miRNAs are known for their ability to down-regulate many cancers as well as for their involvement in cell differentiation and tumorigenesis (Boyerinas et al., 2010). Therefore, LINE-1 is likely to play a principal role in cancer cell regulation and could be a novel breast cancer targeting site.

In vitro studies have shown epithelial cancer cells can be successfully eliminated by some antiretroviral drugs. Evidence suggests that Abacavir and Efavirenz are effective anti-cancer drugs *in vitro* in prostate cancer cell lines (Carlini et al., 2010, Sciamanna et al., 2005). Additionally, Nevirapine is considered to be a potent anticancer drug in terato-, colon-, and

lung carcinomas. Moreover, both Efavirenz and Nevirapine can reduce the growth of melanoma (Houédé et al., 2014). However, previous studies dealing with the repurposing of antiretroviral drugs as anti-cancer drugs have not dealt with TNBCs. Also, there has been little discussion about the signalling pathways affected by antiretroviral drugs in relation to their anti-cancer activity. Hence, in this study, we have focused on repurposing antiretroviral drugs for the treatment of TNBCs, and we have studied potential anti-cancer mechanisms associated with inhibiting LINE-1 expression in TNBCs.

Solid evidence suggests some antiretroviral drugs can effectively eliminate different epithelial cancers in vitro. However, further research needs to be done to prove the effectiveness of these drugs in vivo. One possible explanation is that the drug dosage for treating retrovirus is not sufficient for treating cancers. (Houédé et al., 2014, Hecht et al., 2015). Houédé's group (2018) suggests a sufficient dosage for treating metastatic Castration-resistant prostate cancer could be around 1,200-3,000 mg/day (current dosage in clinical trial: 600 mg/day); Hecht's group (2015) suggests that only 1.5% of patients in clinical trial have received sufficient drugs doses above the in vitro toxic EC50 of Efavirenz. Alternatively, the microenvironment and gene regulation are complicated in the human body. In addition, the existence of CSCs can contribute to drug resistance, thereby reducing the anti-cancer effects of certain antiretroviral drugs.

To date, there has been no study focusing on CSC elimination by antiretroviral drugs. If antiretroviral drugs are only able to eradicate the non-CSCs and not the CSCs, it is likely that cancers will relapse rapidly. Thus, it is critical to investigate whether antiretroviral drugs can eliminate both classes of CSCs. If antiretroviral drugs are indeed effective in treating CSCs, this class of drug should be considered as a valid anticancer treatment aimed at improving breast cancer prognostic outcomes and prolong cancer-free survival for the patients.

1.6 Hypotheses

The main purpose of this study is to examine whether antiretroviral drugs can be repurposed as anticancer drugs to be used against TNBCs. Antiretroviral drugs have been used for treating HIV infections for decades, and recent evidence suggests that they have potential as anticancer drugs given their effectiveness at eliminating certain types of cancer cell (Houédé et al., 2014). Recently, studies of anti-breast cancer effects of antiretroviral drugs have been limited to luminal-type breast cancers including T47D and MCF7 (luminal A) cell lines; however, very little anticancer effects of antiretroviral drugs have been observed in TNBCs which are viewed as the worst type of breast cancers (Patnala et al., 2014). Heterogeneity among breast cancer subtypes may explain why they respond differently to different drugs and different drug concentrations. Additionally, previous studies have failed to consider the effects of drugs on CSCs which are linked to many unfavourable prognostic outcomes. Therefore, this study focuses on the responses of TNBCs and BCSCs to antiretroviral drugs.

The major aim of this study is to explore the efficacy of antiretroviral drugs and their underlying mechanisms for causing the blockade of LINE-1-RT in a range of breast cancers, particularly TNBCs. The other aim of this study is to investigate whether antiretroviral drugs can eliminate CSCs and reduce cancer metastasis and recurrence rates.

In summary, the hypotheses of this study are:

- LINE-1 plays a key role in tumorigenesis in TNBCs.
- Antiretroviral drugs can effectively block LINE-1 expression in TNBCs.
- Antiretroviral drugs can treat TNBCs.
- Antiretroviral drugs can reduce the numbers of CSCs in TNBCs.

To the best of my knowledge, this is the first study to investigate the anticancer effects of antiretroviral drugs in TNBCs. A systematic understanding of how antiretroviral drugs contribute to the anticancer processes in TNBCs is essential for further improving TNBC treatment. Moreover, a thorough examination of the role of LINE-1 in the context of TNBCs could potentially contribute to the field of cancer drug development and treatment. An improved

understanding of the possible links between LINE-1 inhibition and its anti-cancer effects in TNBCs may help us to improve cancer early detection and aid in the valid selection of treatments. It is also anticipated that this research will provide an important opportunity to advance our understanding of CSCs.

1.7 Thesis overview

This chapter briefly summarised the background relating to breast cancers, transposons, and antiretroviral drugs, and it highlighted some of the key questions in the field.

The second chapter describes the materials and methods used in this project.

The third chapter examines whether antiretroviral drugs can affect TNBC cell growth and alter their phenotype. In addition, it investigates whether LINE-1 is involved in the drugs' mechanism of action.

The fourth chapter investigates the role of gene regulation in antiretroviral drug activity.

The fifth chapter tests whether antiretroviral drugs can also be effective at eliminating CSCs and thus reduce the potential for metastasis and patient relapse.

The last chapter analyses and discusses the results, as well as the limitations of our findings. Finally, it examines how these discoveries can improve our knowledge of cancer biology.

Chapter 2 Materials and Methods

2.1 Materials

2.1.1 Chemicals

All chemicals used in this thesis project were of the highest analytical grade unless otherwise specified.

2.1.2 Buffers and solutions

Table 2-1 Buffers and solutions being used in this thesis project

Buffer name	Formula
Phosphate buffered saline (PBS) buffer (1X)	0.137 M NaCl/0.0027 M KCl/0.01 M Na_2HPO_4/0.0018 M KH_2PO_4 in Milli-Q™ water, pH 7.4
Tris-buffered saline (TBS) buffer (1X)	50 mM Tris-Cl/150 mM NaCl in Milli-Q™ water, pH 7.5
TBST	0.05% Tween 20 in TBS
LB broth medium	1% Bacto-tryptone/0.5% Yeast extract/1% NaCl in Milli-Q™ water, pH 7.0
WB transfer buffer	25 mM Tris-base/192 mM Glycine/20% Methanol in Milli-Q™ water
WB blocking buffer	5% non-fat milk or 3% bovine serum albumin (BSA) in TBST
IF fixation solution	2-4% Paraformaldehyde dissolved in PBS
IF blocking buffer	5% BSA/0.3% Triton™ X-100 in PBS
IF antibody dilution buffer	1% BSA/0.3% Triton™X-100 in PBS
PI staining solution	20 μg/ml PI + 0.1% Triton-X 100 in DPBS
Annexin-binding buffer	10 mM HEPES/140 mM NaCl/2.5 mM $CaCl_2$ in Milli-Q™ water, pH 7.4
Crystal violet solution	0.5% Crystal violet in 25% Methanol

(WB: western blot; IF: immunofluorescence; PI: propidium iodide)

2.1.3 Antibiotics

Table 2-2 Antibiotics being used in this thesis project

Antibiotic name	Brand	Cat#
Ampicillin	Sigma-Aldrich	A1593
Kanamycin sulfate	Sigma-Aldrich	60615
Penicillin -Streptomycin	Gibco™	15140122
Geneticin™ (G418)	Gibco™	10131035
Puromycin	InvivoGen	ant-pr

2.1.4 Kits

QIAGEN plasmid Mini and Midi Purification Kits (QIAGEN)

PureLink® RNA Mini Kit (Invitrogen™)

*mir*Vana™ miRNA Isolation Kit (Invitrogen™)

TURBO DNA-*free*™ Kit (Invitrogen™)

SuperScript® IV First-Strand cDNA Synthesis Reaction (Invitrogen™)

PowerUp™ SYBR® Green PCR Master Mix (2X) (Applied Biosystems)

Pierce™ BCA Protein Assay Kit (Thermo Scientific™)

Pierce™ ECL Western Blotting Substrate (Thermo Scientific™)

Cell Line Nucleofector® Kit V (Lonza)

TaqMan® Advanced miRNA cDNA Synthesis Kit (Applied Biosystems™)

TaqMan® Fast Advanced Master Mix (2x) (Applied Biosystems™)

TaqMan® Advanced MicroRNA Assays (Applied Biosystems™)

XTT Cell Viability Kit (Cell Signaling)

AlamarBlue™ Cell Viability Assay (Invitrogen™)

CellTrace™ CFSE Cell Proliferation Kit (Invitrogen™)

FITC Annexin V/Dead Cell Apoptosis Kit with FITC Annexin V and PI, for Flow Cytometry (Invitrogen™)

NuPage™ 4%-12% Bis-Tris Protein Gel Kit (Invitrogen™)

ALDEFLUOR™ Kit (STEMCELL Technologies)

MammoCult™ Human Medium Kit (STEMCELL Technologies)

2.1.5 Antibodies

All the antibodies being used in this study (Table 2-3) are commercially available except anti-human-(LINE-1) ORF1p and anti-human-(LINE-1) ORF2p. These two antibodies were generated in rabbits by IMVS Pathology, Australia.

Table 2-3 Antibodies being used in this thesis project

Name	Brand	Cat#	Application	Dilution
Anti-DDX5	Abcam	Ab21696	WB	1:1000
Anti-Human Vimentin antibody	BD	AF2105	WB	1:800
Purified Mouse anti-CDC42	BD	610929	WB	1:500
Anti-*phospho*-p53 Ser15	Invitrogen	700439	WB	1:200
Anti-HIV-1 RT	Fitzgerald	20-000511	WB	1:1000
Anti-Human-(LINE-1) ORF1p	IMVS	AB3412	WB	1:1500
Anti-Human-(LINE-1) ORF2p	IMVS	AB4213	WB	1:2000
Anti-β-actin	Sigma	A5441	WB	1:10000
Anti-E-cadherin	CST	3195S	WB, IF	1:800, 1:200
Anti-*p*H2A.X (Ser139)	Millipore	05-636-I	WB, IF	1:500, 1:200
Anti-SNAI2/SLUG	BD	564614	WB, IF	1:1000, 1:250
Anti-α-tubulin antibody	Sigma	T9026	WB, IF	1:10000, 1:500
Goat anti-Rat IgG H&L HRP	Abcam	Ab97057	WB	1:2000-1:10000
Goat anti-Mouse IgG (H+L) HRP	Invitrogen	626520	WB	1:2000-1:10000
Goat anti-Rabbit IgG (H+L) HRP	Invitrogen	656120	WB	1:2000-1:10000
Anti-CD24 antibody (FITC)	Abcam	Ab30350	IF	1:50
PE-Cy7™ Anti-CD24	BD	561646	FC	1:50
Alexa Fluor® 647 anti-CD24	BD	561644	FC	1:50
Alexa Fluor® 647 anti-Ki-67	BD	561126	FC	1:50
Alexa Fluor® 488 anti-Annexin V	Invitrogen	V13241	FC	1:20
CD133/1 (AC133)-APC	Miltenyi Biotec	130-090-827	FC	1:20
FITC-Goat anti-Rabbit IgG (H+L)	Invitrogen	65-6111	IF	1:500
Cy5 Goat anti-mouse IgG (H+L)	Invitrogen	A10524	IF	1:500
Rat IgG (Alexa Fluor 647)	Abcam	Ab150159	IF	1:500
Anti-Mouse Alexa 488	CST	4408	IF	1:500
Anti-Rabbit Alexa 555	CST	4413	IF	1:500

(WB: western blot; IF: immunofluorescence; FC: flow cytometry)

2.1.6 Oligonucleotides

The primers being used in the general quantitative real-time polymerase chain reaction (qRT-PCR) were all synthesised by Sigma-Aldrich Pty Ltd. The primer sequences were listed below (Table 2-4).

Table 2-4 Primers being used in qRT-PCR

Name	Sequence	Tm°C	%GC	size	product
PUM1-2-F	CATGGGTCCAGATTCATTCAG	59.4	47.6	21 mer	82 mer
PUM1-2-R	CCTGGAGGATTTCATTGAAGAC	60.3	45.5	22 mer	
ORF1-1-F	TCAAAGGAAAGCCCATCAGACT	61.1	43.5	22 mer	74 mer
ORF1-1-R	TTGGCCCCCACTCTCTTCT	58.4	61.1	19 mer	
ORF2-EN-F	CCCATCAGTGTGCTGTATTC	58.4	50	20 mer	87 mer
ORF2-EN-R	TGGTAGATCTTCCTCCATCC	58.4	50	20 mer	
ORF2-RT-F	ACTGGAAGCATTCCCTTTGAA	57.4	42.9	21 mer	89 mer
ORF2-RT-R	TGGCCAGAACTTCCAACACT	58.4	50	20 mer	

(Tm: m*elting temperature*)

The TaqMan™ Advanced MicroRNA assays being used in the microRNA qRT-PCR were all purchased from Thermo Fisher Scientific. The target microRNA sequences were listed below (Table 2-5).

Table 2-5 Mature miRNA sequence of the target miRNAs being used in microRNA qRT-PCR

Target microRNA	Assay ID	Mature miRNA sequence
miR-21	rno481342_mir	UAGCUUAUCAGACUGAUGUUGA
miR-182	477935_mir	UUUGGCAAUGGUAGAACUCACACU
let-7a	478575_mir	UGAGGUAGUAGGUUGUAUAGUU
miR-423	478090_mir	UGAGGGGCAGAGAGCGAGACUUU

2.1.7 Plasmid DNA constructs

2.1.7.1 pmaxGFP® control plasmid DNA for Nucleofection®

pmaxGFP® (Figure 2-1) which is a positive control plasmid of Nucleofection® is provided in the Amaxa® Cell Line Nucleofector® V Kit (Lonza). This plasmid can express Green Fluorescent Protein (GFP, from *Pontellina p.*) in successfully transfected cells. Transfection efficiency can vary depending on different batches of experiments. Therefore the transfection efficiency of Nucleofection® should be analysed and monitored by observing the pmaxGFP transfected control cells in each experiment.

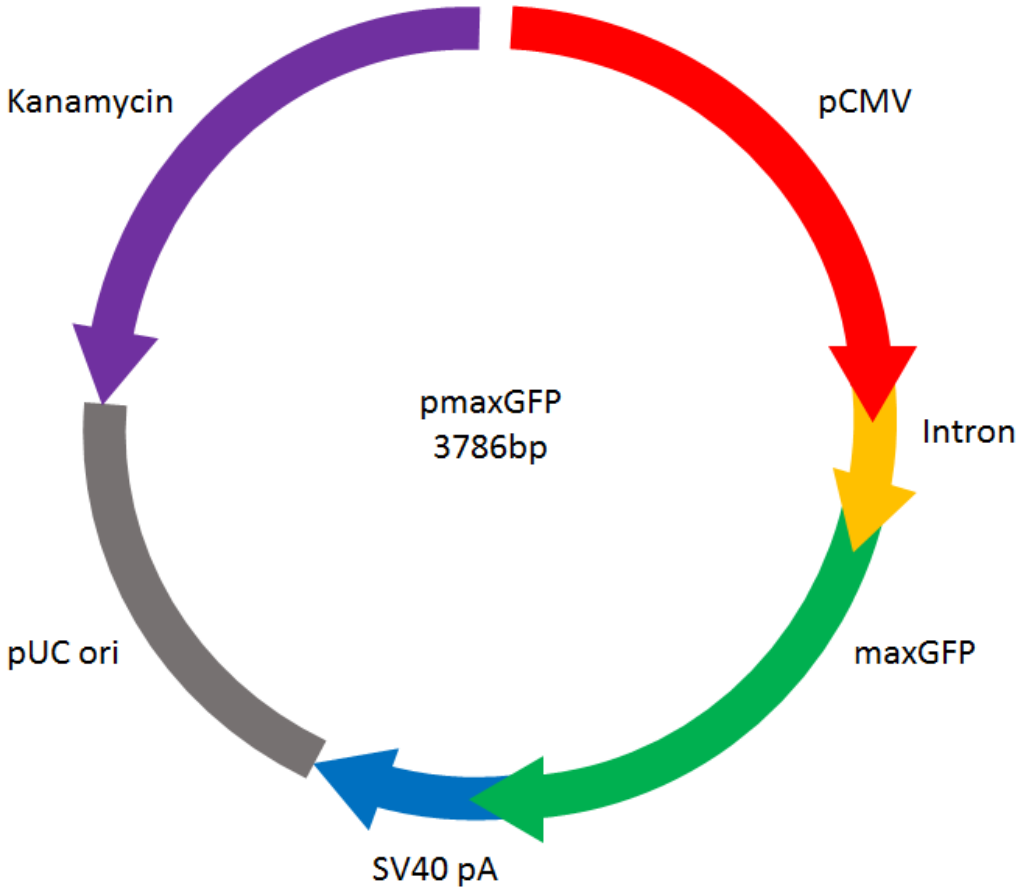


Figure 2-1 The plasmid DNA map of pmaxGFP positive control plasmid of Nucleofection®
(Referring from Olmedillas López et al. (2016))

2.1.7.2 pBS-L1-PA1-mneo plasmid DNA

pBS-L1-PA1-mneo plasmid was used for simulating retrotransposition activity of LINE-1 in mammalian cells (Wagstaff et al., 2011). This plasmid contains an optimised LINE-1 sequence (with ~75% similarity to the original LINE-1 sequence) driven by the CMV promoter and a neol indicator (Figure 2-2). The neol indicator includes a reversed neomycin resistant sequence without any function because of intron disruption, only if retrotransposition happens due to the function of LINE-1, there is a chance that a functional neomycin resistance gene which is driven by the SV40 promoter can be expressed. With the neomycin resistant ability, the transfected cells can survive in the Geneticin-containing medium (Geneticin is a neomycin analogue for treating mammalian cells).

Figure 2-2 The plasmid DNA map of pBS-L1-PA1-mneo plasmid.
The plasmid map is created by SnapGene® Viewer, the sequence is cited from Addgene (n.d.) *pBS-L1-PA1-mneo (Plasmid #51288)*. Retrieved from: https://www.addgene.org/51288/

2.1.7.3 pSM2 and pUTR plasmid DNA

Arrest™ pSM2 control vector (OpenBiosystems, Cat# RHS1704, Figure 2-3), a pSHAG-MAGIC2 plasmid without short hairpin RNA (shRNA) insertion is a negative control of any shRNA expressed by the Expression Arrest™ shRNA system (Paddison et al., 2004). The shRNA insertion site of pSM2 can be replaced by a designed shRNA, and this designed shRNA can adapt a hairpin formation that degrades later to small interfering RNA (siRNA). The siRNA expresses its function of RNA interference, inhibiting target gene expression (Paddison et al., 2004). The Expression Arrest™ shRNA system can be an effective tool for knocking-down target genes. pUTR is a previously designed LINE-1 inhibiting plasmid utilising the Expression Arrest™ shRNA system and can specifically target the LINE-1 promoter. The sense and antisense sequences of the Expression Arrest™ shRNA plasmids used in this thesis are listed in Table 2-6.

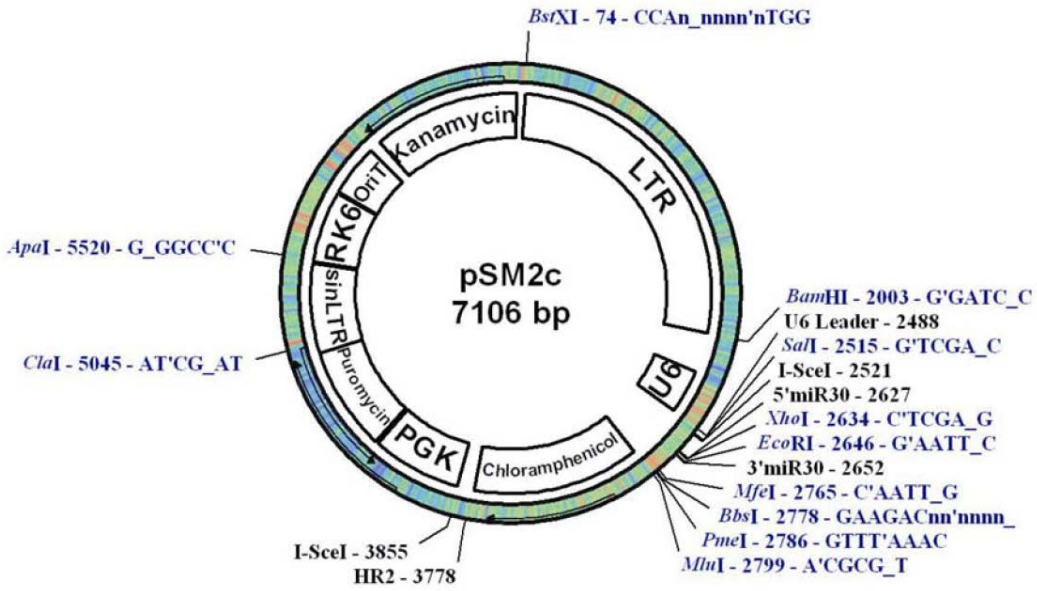


Figure 2-3 The plasmid DNA map of Arrest™ pSM2 empty vector
Cited from: OpenBiosystems (n.d.) *Expression Arrest™ pSM2 empty vector*. Retrieved from http://www.pages.drexel.edu/~bio/cores/Documents/Protocols/(Prot%204).pdf

Table 2-6 The sequences of shRNAs in pSM2 vector shRNA insertion site

pSM2	
Sense (mir5')	UUCUCCGAACGUGUCACGUdTdT
Anti-sense (mir 3')	ACGUGACACGUUCGGAGAAdTdT
pUTR	
Sense (mir 5')	TTTACCTAAGCAAGCCTGGGC
Anti-sense (mir 3')	TGCCCAGGCTTGCTTAGGTAAA

2.1.8 Cell lines

The cell lines used in this thesis project include two luminal types of breast cancer cell lines – T47D, MCF7, six triple-negative breast cancer (TNBC) cell lines – MCF10AT, MCF10CA1α, MDA-MB-231, MDA-MB-436, BT-549, BT-20, and a non-tumorigenic epithelial cell line – MCF10A; all of which are listed in Table 2-7 with their classifications. All cell lines mentioned were purchased from the American Type Culture Collection (ATCC) unless otherwise specified.

MCF10A, MCF10AT, and MCF10CA1α are a breast cancer model system that represents breast cancer progression within these cell lines. (Lim et al., 2015). MCF10A is a non-cancerous control cell line; MCF10AT is a premalignant cell line which is derived from MCF10A; and MCF10CA1α is a fully malignant cell line which is derived from MCF10AT (Imbalzano et al., 2009). With these cell lines, the net influence of cancer development can be studied since these cell lines are from the same origin and therefore have a similar genetic background.

Table 2-7 Cell lines being used in this thesis project

Cell line	Receptors	Molecular type	TN subtype	Pathology
MCF10A	$ER^-/PR^-/HER2^-$	TN	N/A	Fibrocystic disease
MCF10AT	$ER^-/PR^-/HER2^-$	TN	Unknown	-
MCF10CA1α	$ER^-/PR^-/HER2^-$	TN	Unknown	-
MDA-MB-231	$ER^-/PR^-/HER2^-$	TN	MSL	AC
MDA-MB-436	$ER^-/PR^-/HER2^-$	TN	MSL	AC
BT-549	$ER^-/PR^-/HER2^-$	TN	M	IDC
BT-20	$ER^-/PR^-/HER2^-$	TN	Unclassified	IDC
T47D	$ER^+/PR^+/HER2^-$	Luminal A	N/A	IDC
MCF7	$ER^+/PR^+/HER2^-$	Luminal A	N/A	IDC

(ER: Estrogen receptor; PR: progesterone receptor; HER: human epidermal growth factor receptor 2; TN: triple-negative; M: *mesenchymal-like*; MSL: *mesenchymal stem-like;* AC: adenocarcinoma; IDC: invasive ductal carcinoma; N/A: not applicable) (Holliday and Speirs, 2011, Dai et al., 2017)

2.1.9 Other materials

All the schematic diagrams were created by BioRender online software (https://biorender.com/) unless otherwise specified.

2.2 Methods

2.2.1 Mammalian cell culture

2.2.1.1 Traditional two-dimensional cells

MCF10A, MCF10AT, and MCF10CA1α cultured in the complete DMEM/F-12 medium (Gibco™, Cat# 10565018) with 5% horse serum (Gibco™, Cat# 16050122), 10 µg/ml Insulin (Sigma-Aldrich, Cat# I882), 20 ng/ml Epidermal Growth Factor (Sigma-Aldrich, Cat# E4127), 0.5 µg/ml Hydrocortisone (Sigma-Aldrich, Cat# H0888), and 100 ng/ml Cholera toxin (List Biological Laboratories, Cat# 101B). T47D, MCF7, MDA-MB-231, MDA-MB-436, and BT-549 cells were cultured in the complete DMEM medium (Gibco™, Cat# 11965118) with 10% foetal bovine serum (FBS) (Thermo Scientific, Cat# SH30071.03). And BT-20 was cultured in the complete RPMI 1640 medium (Gibco™, Cat# 61870127) with 10% FBS. All cell lines were cultured in a 37°C incubator with 5% CO_2 supply. Cells were subcultured every two to four days contingent upon cell proliferation rates and experimental designs. Cell cultures were regularly observed and monitored under a CK2 Microscope (Olympus). A haemocytometer and a microscope were employed for calculating cell number as needed. Most of the cell culture experiments in this study were performed under traditional two-dimensional cell culture conditions unless otherwise specified.

2.2.1.2 MammoCult™ tumorsphere formation three-dimensional cell culture

MammoCult™ Human Medium Kit (STEMCELL technologies, Cat# 05620) had been utilised in many breast cancer cell lines and primary breast cancer tissues for generating tumorspheres and have been recognised to represent the cells with stem cell-like characteristics. After harvesting and counting cells from a traditional two-dimensional culture, $2x10^4$ cells were seeded into an individual well of an Ultra-Low Adherent Suspension Culture plate (Corning, Cat# 3471) and were then cultured in the completed serum-free MammoCult™ medium with 4 µg/mL Heparin (STEMCELL technologies, Cat# 07980) and 0.48 µg/mL Hydrocortisone for three to seven days in order to observe tumorsphere formation with/without drug treatment (Figure 2-4). The number and size of tumorspheres in each well were measured for further analysis.

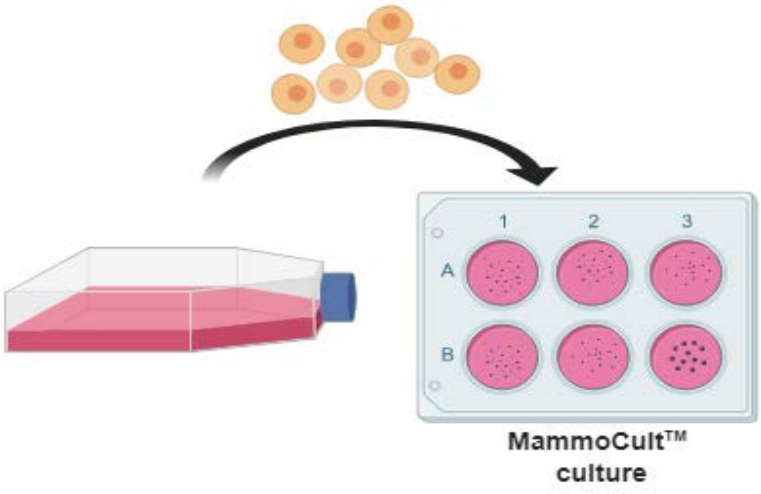


Figure 2-4 MammoCult™ tumorsphere formation
Harvested cells from 2D culture are seeded in a 6-well plate and cultured by completed MammoCult™ medium. After three to five days, some tumorspheres can be formed depending upon the different characteristics of cell lines. These tumorspheres are observed by an Olympus IX 81 microscope, and the number and size are measured by Image J software.

2.2.1.3 Drug treatment

Tested drugs were all dissolved in dimethyl sulfoxide (DMSO) to make 100 mM stock and were stored in -20°C for long term storage. When performing drug treatment experiments, appropriate concentrations of drugs (EC50 value for each cell line) were mixed with the cell culture medium, and the pre-seeded cells were incubated with the drug-contained medium for a certain period of time (four days for the traditional two-dimensional culture; seven days for the MammoCult™ tumorsphere three-dimensional culture) in a 37°C incubator with 5% CO_2 supply. The drug-contained medium was replaced by the fresh drug-contained medium every 48 hours unless otherwise specified.

2.2.2 Mammalian cell transfection

2.2.2.1 Plasmid amplification and purification

The frozen glycerol stocks of *E. coli* containing target plasmids were firstly cultured for 16 hours in 5 ml LB broth medium (mentioned in section 2.1.2.) with an appropriate amount of antibiotics. 100 µl of the primary-cultured *E. coli* were then cultured in 200 ml LB broth medium with an appropriate amount of antibiotics for four to eight hours depending on the speed of bacteria growth. The plasmids within *E. coli* were amplified with bacterial amplification and were further isolated by a Plasmid Purification Kit (QIAGEN, Cat# 12943) following the given protocol. After collecting cell suspension of the cultured *E.coli* in the centrifugal vessel, the bacterial cells were spun down by centrifugation. The bacterial pellet was then resuspended in 4 ml of P1 buffer. 4 ml of P2 buffer was also mixed with the mixture followed by gentle mixing of the vessel. An extra 4 ml of P3 buffer was mixed with bacteria-P1-P2-mixture and further incubated on ice for 15 minutes. The sample was then spun down at 4°C for at least 30 minutes to remove the residual precipitate. The supernatant containing target plasmids was passed through a column which was designed to bind plasmid DNA. Later, 3.5 ml of isopropanol and 70% ethanol was added to the column orderly so as to precipitate plasmid DNA. The final DNA precipitate was resuspended in 50-100 µl of autoclaved water for further experiments.

2.2.2.2 Nucleofection®

Nucleofection® can be a powerful tool for cell transfection. All the cell transfection experiments in this study were conducted by Nucleofection®. Nucleofector® I machine (Lonza) and Cell Line Nucleofection® Kit V (Lonza, Cat# VVCA-1003) were utilised for transfecting plasmid DNA into certain types of mammalian cells. First, target cells were cultured for two to three days to approximately 80% confluency. The cells were then harvested and approximately a million cells were collected and mixed with 100 µl Nucleofector® reagent and 5 µl plasmid DNA. The mixture was transferred into a Nucleofector® cuvette, which was then placed into a Nucleofector® I machine. After applying for a suitable Nucleofection® programme with the cells based on previous experiments (X-05 program for T47D cell lines; T-20 program for MCF10A, MCF10AT, and MCF10CA1α), the operated cells were transferred into the pre-warmed cell culture medium. 24 to 48 hours later, the normal cell culture medium was exchanged with the transfection selection medium, which contains antibiotics such as

Puromycin and Geneticin™, to eliminate non-transfected cells and allowing only successfully transfected cells to survive.

2.2.3 Protein detection

2.2.3.1 Western blotting

Western blotting is a commonly used method for detecting relative protein expression among samples. The protocol of western blotting used in this project was described as following. Samples were harvested from tissue culture flasks and then were lysed by using a Radioimmunoprecipitation (RIPA) assay protein extraction buffer (Sigma, Cat# R0278) laud with an appropriate amount of protease inhibitors (Roche, Cat# 11836170001) for preventing protein digestion. Afterwards, proteins in the samples were separated according to their molecular weights through a NuPAGE 4-12% Bis-Tris Protein Gel (Thermo Fisher Scientific, Cat# NP0321), followed by transfer to a pre-soaked nitrocellulose membrane (Bio-rad, Cat# 14024) within a Bio-Rad Mini Trans-Blot® system (Bio-rad, Cat# 1703930) containing WB transfer buffer. A protein ladder was used (Novex Sharp Pre-stained protein standard (Thermo Fisher Scientific, Cat# LC5800)). Subsequently, the transferred membrane was incubated with targeted primary antibodies such as anti-HIV-1 RT antibodies (for detecting ORF2p), anti-ORF1, anti-β-actin, anti-α-tubulin (as a loading control), anti-fibronectin, anti-E-cadherin, anti-SLUG, or other relevant primary antibodies diluted in the WB blocking buffer. This process was followed by incubating appropriate types and amount of secondary antibodies such as anti-Rabbit-HRP (horseradish peroxidase), anti-Rat-HRP, anti-Goat-HRP, and anti-Mouse-HRP. The details of these antibodies were listed in Table 2-3. After incubating the stained membranes with enhanced chemiluminescent substrates (ECL) (ThermoFisher, Cat# 32109), these stained membranes were exposed in a GE LAS4000 Fluoimager for detecting relative HRP activities among samples. Results were exported from the software for further analysis.

2.2.3.2 Immunofluorescence (IF)

Immunofluorescence (IF) is a common method for detecting protein expression patterns within cells or on the cell surface. Several different cell markers including CD24, E-cadherin, Fibronectin, and SLUG were investigated by IF experiments in order to understand possible cell molecular and morphological changes under alternative treatment conditions. These markers were directly targeted by specific antibodies which were listed in section 2.1.3. The protocol of IF being used in this project is described below (Figure 2-5).

Cells were seeded on a 12 mm Poly-D-lysine coating glass coverslip (for enhancing cell adhesion to glass surface) (Neuvitro, Cat# GG-12-1.5-PDL) in a 24-well plate for one day until the cells attached to the coverslip surface. These cells were then treated with drugs or DMSO (control) for another four days. Before staining, cells were washed with PBS for three times, and these cells were fixed on coverslips by IF fixation solution (mentioned in section 2.1.2) for 15 minutes. Afterwards, the cells were permeabilised by 0.25% Triton X-100 for 10 minutes followed by incubating with IF blocking buffer for at least 1 hour. The permeabilised cells then were incubated with 1 in 200 to 500 IF antibody dilution buffer diluted primary antibodies for 16 hours in a 4°C cold room with continuously gentle rotation. Subsequently, the cells were washed with PBS followed by incubating with 1 in 500 IF antibody dilution buffer diluted secondary antibodies if applicable (some primary antibodies are conjugated with fluorescent dyes, secondary antibodies are unnecessary for them). Finally, the coverslip with stained cells was mounted with ProLong™ Gold /Diamond Antifade Mountant with DAPI (nucleus indicator) (Molecular Probes®, Cat# P36930/P36961). The coverslips were observed under a Leica SP5 confocal microscope after 24 hours incubation with the mounting buffer in room temperature protecting from lights. The fluorescent signals of the IF samples can be maintained for several months.

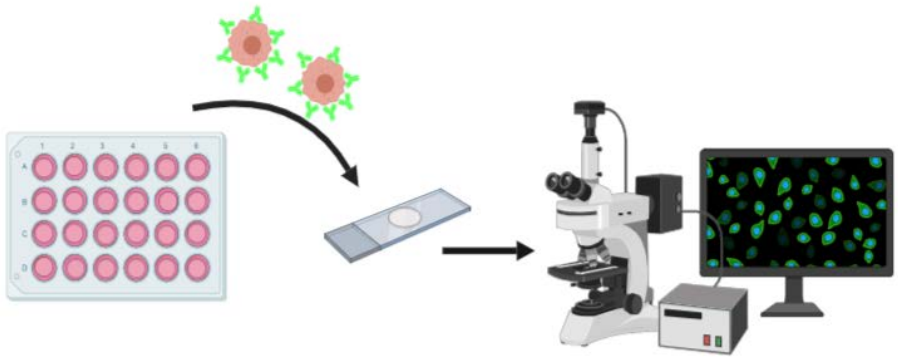


Figure 2-5 Schematic diagram for the process of immunofluorescence
Cells are seeded on a coverslip which is covered with the complete culture medium. These cells are treated by the drugs once they attached to the coverslip. After a treating period, cells on the coverslip are stained with fluorescent antibodies which can target specific cell markers. The stained samples are mounted on a microscope slide and are observed under a confocal microscope.

2.2.4 RNA detection

2.2.4.1 RNA isolation

Total RNA extraction was basically based on the protocol of PureLink® RNA mini Kit including the following steps:

Lysis and homogenisation: Samples were washed with PBS, and then lysed in lysis buffer (2-mercaptoethanol was added freshly before use). After collecting all of the lysates, a homogeniser was used with RNase-free autoclaved tips to homogenise the lysate for at least 1 minute. The homogenised lysate then was centrifuged at 2600 xg for 5 minutes, and the supernatant was transferred to a new clean tube.

Binding and washing: Equal amount of 70% ethanol was mixed with cell lysate, and the mixture was passed through a spin column by centrifuge. After that, the column was washed with Wash Buffer I by centrifuge.

Removing DNA and elution: The column was incubated with 60 µl of TURBO DNA-free (Life Technologies) mixture (6 µl 10x DNase Buffer and 4 µl DNase diluted in 50 µl Nuclease-free water) at room temperature for 30 minutes, and then the column was consequently washed with Wash Buffer I and II. Finally, RNA was eluted with nuclease-free water and the quality of the RNA was assessed using an Aglient 2100 bioanalyzer (Aglient technologies).

The method used for extracting microRNA was similar to the total RNA extraction method, but a different kit (*mir*Vana™ miRNA isolation kit) was utilised as a pre-requisite requirement for TaqMan® Advanced microRNA assays used in microRNA experiments. Basically, a Phenol: Chloroform extraction step was added in between homogenisation and RNA binding steps, and the flow-through was collected in the first filter step in order to accumulate microRNA and reduce the influence from mRNA. After complete washing and elution, isolated miRNAs were then quantified using a Nanodrop.

2.2.4.2 Complementary DNA (cDNA) synthesis

SuperScript® IV First-Strand cDNA Synthesis Reaction was used for synthesising cDNA from target RNA. In this assay, the SuperScript® IV was mainly used for general Quantitative real-time PCR (qRT-PCR) in LINE-1 expression experiments, the protocol was as follows. For each sample, 5 µg RNA template was mixed with 1 µl Random hexamers (50 µM), and 1µl dNTP mix (10 mM), and made up to a total volume of 13 µl by adding DEPC-treated water in an Eppendorf

tube. The mixtures were incubated at 65°C for 5 minutes for RT primer annealing and incubated on ice for 1 minute. Afterwards, the annealed RNA was mixed with 4 µl 5x SSIV buffer, 1 µl 100 mM DTT, 1 µl RNaseOUT™ RNase Inhibitor, and 1 µl SuperScript® IV RT (200 U/µl). This reaction mixture was incubated at 23°C for 10 minutes for reverse transcription and cDNA synthesis, followed by an inactivation step at 80°C incubation for 10 minutes. The RNA-cDNA mixture was then incubated with 1 µl RNase H (2 U/µl) at 37°C for 20 minutes for reducing RNA contamination for the following experiments. The synthesised cDNA was stored at -20°C until further use or its concentration determined by a NanoDrop™ machine (Thermo Scientific™) for directly conducting qRT-PCR experiments.

TaqMan® cDNA synthesis assay was used for microRNA studies. This assay can be separated into four major parts including the polyadenylation (poly (A)) tailing reaction, the adaptor ligation reaction, the RT reaction, and the miR-Amp reaction. For elongating poly (A) tail to microRNA, 10 ng of total RNA (made up to 2 µl) was mixed with 0.5 µl poly (A) buffer (10x), 0.5 µl ATP, and 0.3 µl poly (A) Enzyme in a tube, and the total volume was made up to 5 µl. The mixture was then incubated at 37°C for 45 minutes for polyadenylation, followed by a reaction termination step at 65°C for 10 minutes. The poly (A) reaction product was then incubated with 3 µl DNA Ligase buffer (5x), 4.5 µl 50% PEG 8000, 0.6 µl ligation adaptor (25x), 1.5 µl RNA ligase, and 0.4 µl RNase-free water at 16°C for 1 hour in order to complete the ligation step. Afterwards, the adaptor ligation reaction product was mixed with RT reaction mix containing 6 µl RT buffer (5x), 1.2 µl dNTP mix (25 mM), 1.5 µl universal RT primer (20x), 3 µl RT enzyme mix (10x), and 3.3 µl RNase-free water. The mixture was incubated at 42°C for 15 minutes followed by 85°C for 5 minutes for completing reverse transcription. The RT reaction product needs a further step, miR-Amp reaction, before performing the real-time PCR (RT-PCR). This reaction can uniformly increase the amount of cDNA for miRNAs. 5 µl of RT reaction product was mixed with 25 µl miR-Amp Master Mix (2x), 2.5 µl miR-Amp Primer Mix (20x), and 17.5 µl RNase-free water. This mixture was placed into a PCR machine, and the cDNA was amplified by following steps: an enzyme activation step (95° for 5 minutes), 14 cycles of denature/anneal/extend cycle steps (95°C for three seconds followed by 60°C for 30 seconds), and a stop reaction step (99°C for 10 minutes). The product then can be directly used for further RT-PCR experiments or be stored at -20°C for long term storage.

2.2.4.3 General quantitative real-time PCR (qRT-PCR)

qRT-PCR is a technique which is widely used for quantifying relative RNA levels between samples. Principally, the more PCR cycles being detected in a sample mean the lower amount of cDNA is contained in this sample, vice versa. The details of the qRT-PCR protocol being used in this project were described below.

96-well qRT-PCR plates (Applied Biosystems) were used in this study. One well contained one sample per reaction. Each reaction included 10 µl Power SYBR Green PCR Master Mix (2x) (Applied Biosystems), 0.5 µl forward primer (100 mM), 0.5 µl reverse primer, and 10 ng corresponding sample cDNA, the total volume was made up to 20 µl. The sample loaded-plate was then covered and sealed with an adhesive cover and was centrifuged to ensure the sample containing the same amount of reagents. The qRT-PCR was operated StepOnePlus™ RT-PCR instrument through accompanying software - the ExpressionSuite™ software. The setting of the thermal program for the qRT-PCR included a 10 minutes enzyme activation step (at 95°C), and 40 PCR cycle (denature at 95°C for 15 seconds and anneal/extend at 60°C for 1 minute). The comparative Ct ($\Delta\Delta C_t$) was used to perform the relative quantification of the qRT-PCR results. These results were then imported to GraphPad Prism 6 for further graphical analysis.

2.2.4.4 microRNA qRT-PCR

The principle of microRNA qRT-PCR is similar to general qRT-PCR, however, because of low copy numbers and a relatively small amount of microRNAs compared with mRNAs, another qRT-PCR system was used with different types of indicators and primers with the general qRT-PCR. TaqMan® Advanced microRNA assay; (Applied Biosystems™) which contains PCR reagents, reaction buffer, and primers are sensitive to small amounts of sample, and it can measure mature microRNAs from less than 10 ng of the total RNA sample.

In this assay, the total RNAs was prepared by *mir*Vana™ miRNA Isolation kit in order to enrich small RNAs. The next step was to synthesise cDNA by the TaqMan® Advanced miRNA cDNA Synthesis Kit. Afterwards, 5 µl of one in ten diluted cDNA templates were mixed with 10 µl TaqMan® Fast Advanced Master Mix (2X), 1 µl TaqMan® Advanced miRNA assay (20X), and 4 µl RNase-free water. The mixture was then transferred into a 96-well PCR plate and measured by the StepOnePlus™ RT-PCR instrument (Applied Biosystems). The PCR reactions were set as one enzyme activation step (95°C, 20 seconds) and 40 cycles of denature step (95°C,

one second) and anneal/extend step (60°C, 20 seconds). The ExpressionSuite™ software was used for analysing the results by comparing the relative quantification $\Delta\Delta C_t$ values among samples. All the samples had four technical replicates to ensure consistency.

2.2.5 Cell viability and cell proliferation assay

2.2.5.1 Crystal violet cell viability assay

Crystal violet (Sigma-Aldrich, Cat# C0775), a dark-purple colour dye, can stain proteins and DNA, thus it can be used as a cell indicator. Crystal violet assay is a quick screening method for indirectly determining relative numbers of adherent cells on the same size of cell culture plates since dead adherent cells detach from the plate but live adherent cells remain on the plate. The violet assay as described below.

The cultured cells in 25 cm^2 flasks are gently washed with water until no red-colour medium can be seen and the remaining liquid is aspirated by filter papers. Afterwards, an appropriate amount of Crystal violet solution (mentioned in section 2.1.2) is added to the flasks and covers the cell-attached surface of the flasks. The solution is incubated with the cells for 20 minutes with gentle shake following by another washing step for excluding redundant crystal violet solution. The washed flasks are then air-dried for 24 hours. Subsequently, 500 µl of methanol is added to the flasks and incubated with the stained purple-colour cells for at least 30 minutes until the cells become transparent. The de-stained methanol buffers are measured by SmartSpec™ 3000 (BIO-RAD) at OD570 nm, and the relative cell viability of each treating conditions can be calculated.

This method is inexpensive and can roughly quantify cell viability of large-scale of adherent cells; however, the accuracy might be lower than other methods which can directly measure enzyme activities within the cells. Thus, additional cell viability assays were used for further studying drug responses of the cells.

2.2.5.2 XTT cell viability assay

2,3-Bis-(2-methoxy-4-nitro-5-sulfophenyl)-2H-tetrazolium-carboxanilide (XTT) cell viability assay (Cell Signaling, Cat# 9095) is a commonly used colourimetric assay. It has been utilised in both cell viability and proliferation experiments. Dissolved XTT is a slightly yellow reagent; whereas, the colour of XTT changed to bright orange when it is reduced by dehydrogenase enzymes. The N-Methylphenazonium methyl sulfate in the XTT solution can enhance XTT reduction, thus the colour change can be violent. These dehydrogenase enzymes lose their functions immediately after cell death; therefore, measuring the intensity of reagent colour can identify the relative numbers of viable cells among treatments.

The XTT assay is conducted by the following steps. After mixing XTT reagent and electron coupling solution (50:1 ratio), 50 µl of the complete XTT detection reagent is added to each testing well of 96-well plates, then the plates are incubated for 4 hours in a 37°C incubator for an enzyme reaction. Finally, the absorbance of each well is measured at 450 nm using an ELISA reader, and then the data is analysed by Prism software. For cell viability experiment, testing cells are treated with different concentrations of drugs (0 µM, 10 µM, 20 µM, 30 µM, 40 µM, and 50 µM, separately) in order to calculate proper 50% effective cytotoxic concentration (EC50) values of the drugs for further experiments; For cell proliferation experiments, cells are treated with DMSO-untreated-control, Efavirenz (EFV) at EC50 concentration, and SPV122 (SPV) at EC50 concentration for 24 hours, 48 hours, 72 hours, and 96 hours to observe changes in cell proliferation between cells with/without drug treatment.

2.2.5.3 AlamarBlue™ Cell Viability Assay

The functions of alamarBlue™ Cell Viability Reagent (Thermo Fisher Scientific, Cat# DAL1100) are similar to the XTT assay which are used for observing cell viability and cell proliferation. The alamarBlue™ cell viability assay could be better than the XTT cell viability assay since it is less toxic to cells. The principle of the alamarBlue™ cell viability assay is that the blue-colour and non-fluorescent resazurin, the major component of alamarBlue™ reagent, can be reduced to resorufin within live cells, and then the red-colour high-fluorescent resorufin can be measured either by absorbance at 570 nm or fluorescence at 590 nm (excitation at 530 nm to 560 nm). The readout of absorbance or fluorescence can be calculated and therefore the relative cell viability of each well can be determined. The major difference between alamarBlue assay and XTT assay is that after incubating with alamarBlue reagent, the cells could still survive and grow. This assay was conducted to partially replace XTT assay in later of this study.

2.2.5.4 CellTrace™ CSFE cell proliferation assay

Unlike XTT and alamarBlue™ assays, CellTrace™ 5(6)-Carboxyfluorescein N-hydroxysuccinimidyl ester (CSFE) Cell Proliferation Assay (Thermo Fisher Scientific, Cat# C34554) is an assay specifically for detecting cell division (Figure 2-6). CSFE is a non-fluorescent cell-permeant compound, thus it can pass through cell membrane when incubated with cells. The acetate group is then cleaved by enzymes in the cell cytoplasm, thus the remnant green-fluorescent carboxyfluorescein is locked within the cell. When the succinimidyl ester group couple with intracellular free amines, the cell can be detected as a green fluorescent cell

for quite a long term by flow cytometry FITC filter. When cell division is occurring, the green fluorescence is divided into two daughter cells, and therefore the brightness of green fluorescence is reduced to half-way of the original brightness. The brightness of CSFE-tracking cells is inversely related to the frequency of cell division and cell proliferation, therefore the brighter cells have relatively lower cell division rate.

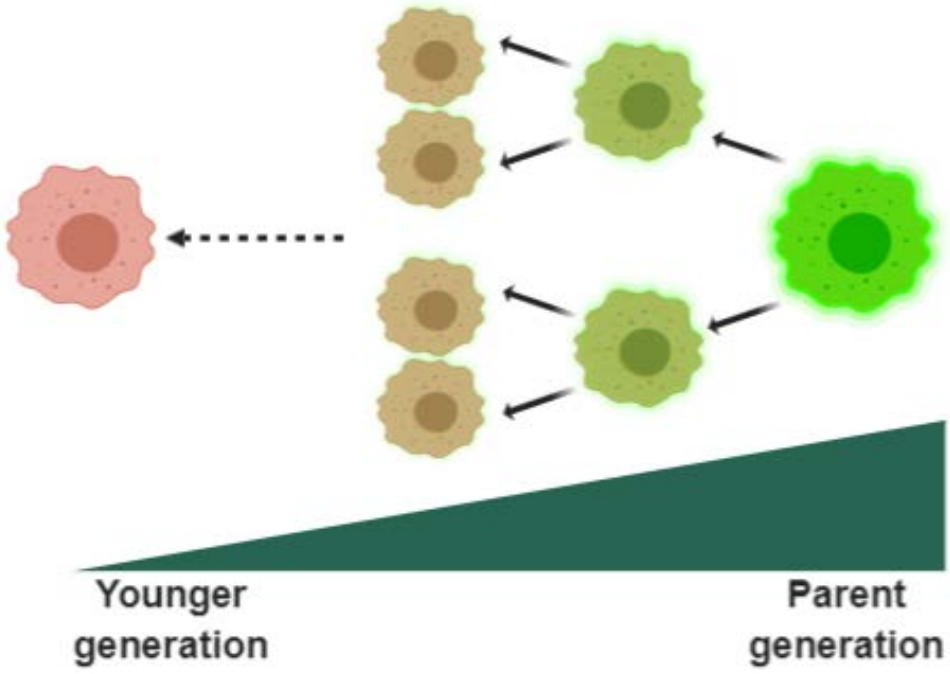


Figure 2-6 Principle of CellTrace™ cell proliferation assay
In CellTrace™ assay, the parent cells are stained with non-toxic CSFE which can be activated its green-fluorescence and be implemented to the cell membrane after enzyme digestion within cells. Therefore, the brightness of green-fluorescence within the cell can only be diluted during cell division. The daughter cells contain about half of the original green-fluorescent compounds. Continuously, the younger generation of cells contain less green-fluorescence. Because of this, the relative cell division rate can be monitored and be compared between different treatments.

2.2.6 Cell physiological and morphological change

2.2.6.1 Cell cycle analysis

Propidium iodide (PI) (Sigma-Aldrich, Cat# P4170), a nucleic acid fluorescent indicator, is commonly used for cell cycle analysis owing to the diverse amount of DNA in different cell stages. DNA variation in cells appears in distinct cell interphase stages of the cell cycle, and this can be detected by PI staining. PI can bind proportionately to the amount of DNA presenting in the cell. In a human cell, DNA synthesis occurs in synthesis (S) phase thus the amount of DNA of S phase cell should be in between of Gap 0 and Gap 1 (G0/G1) phase cell and Gap 2 (G2) phase cell; while the cell contains a double amount of DNA in the G2 phase than those cells in the G0/G1 phase. Therefore, PI staining for the cells in S phase is brighter than the cells in G0/G1 phase but darker than the cells in G2 phase; whereas, the cells in G2 phase are approximately twice as bright as cells in G0/G1 phase. This feature can identify the cells in different cell cycle stages.

This cell cycle experiment was conducted with the following steps. After harvesting drug-treated and untreated cells, the cells were fixed in pre-chilled 70% ethanol prepared in Dulbecco's Phosphate-buffered Saline (DPBS) (Gibco™, Cat# 14190250) for 2 to 24 hours (the fixed cells can be stored for 1 month). The next step was to wash the fixed cells with DPBS for three times, and then to stain the cells with PI staining solution (mentioned in the section 2.1.2) with 0.2 mg/ml RNase A which can reduce RNA contamination in cell cycle experiment for at least 15 minutes. Finally, the stained samples were analysed through flow cytometry Fortessa (BD) at the fluorescence emission at 575 nm to observe the percentages of the cells in different interphase stages.

SubG1 detection is a branch of the cell cycle study. The small fragment of DNA is considered as the products of apoptosis, thus the cell with DNA fragment smaller than a normal cell (in G0 or G1 phase) is defined as an apoptotic cell. Through a fluorescence-activated cell sorting (FACS) analysis, the proportion of apoptotic cell in the entire population can be observed. However, because the subG1 population may include cell debris, the exact apoptotic cells might be less than predicted.

2.2.6.2 Cell apoptosis

One of the most popular methods for detecting apoptotic and necrotic cells is the FITC Annexin V/Dead Cell Apoptosis Kit (Invitrogen™, Cat# V13242). In this study, PI was used to stain DNA, and Annexin V was used for phosphatidylserine staining. PI failed to enter live or early apoptotic cells because of the protection of cell membrane, but PI was able to stain late apoptotic and necrotic cells due to the loss of membrane integrity. Annexin V labelled cells externalised phosphatidylserine on the cell surface in both apoptotic and necrotic cells. Thus, apoptotic cells and necrotic cells could be distinguished. The whole process is listed below: (1) harvested $1x10^6$ cells, (2) stained the cells in 100 µl of Annexin-binding buffer (mentioned in section 2.1.2) with 5 µl of FITC Annexin V solution, 2 µg/ml PI, and 50 µg/ml RNase A in dark environment for 30 minutes, (3) washed the stained cells three times for washing out redundant staining solution, and (4) analysed the stained cells through flow cytometry Fortessa with the fluorescence emission at 530 nm (for detecting FITC signal) and 575 nm (for detecting PI signal) immediately after staining completed. The cells can be separated into three major groups – live cells (Annexin V^-/PI^-), early apoptotic cells (Annexin V^+/PI^-), and late apoptotic and necrotic cells (Annexin V^+/PI^+) according to the FACS analysis results (Figure 2-7).

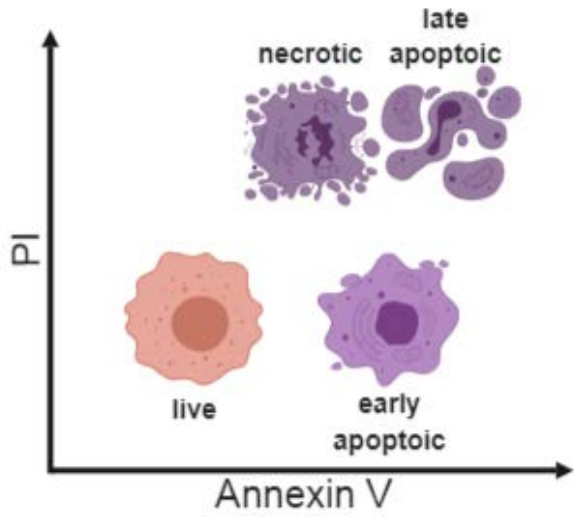


Figure 2-7 Three major cell status being identified by Annexin V/ PI staining
Based on Annexin V/PI staining, cells in different status can be separated by flow cytometry according to their fluorescent patterns. The cells located in Annexin V^-/PI^- region are live cells; the cells located in Annexin V^+/PI^- region are early apoptotic cells; the cells located in Annexin V^+/PI^+ region are late apoptotic and necrotic cells.

2.2.6.3 Cell morphology detection

Cell morphology detection experiments were conducted by Phalloidin staining. Phalloidin, a natural toxin from Amanita Phalloides, has high-affinity to eukaryote-enriched filamentous actin (F-actin), which is an essential element for cytoskeleton. With Phalloidin staining, the entire cell shape and stress fibres within the cell can be clearly observed. Phalloidin staining in this experiment was a modified IF staining, the details of are described below.

Firstly, cells were seeded on 12 mm Poly-D-lysine coating glass coverslips in a 24-well plate for one day until the cells attached to the coverslip surface and then treated these cells with the drugs or DMSO-control for four days. Before staining, cells were washed with PBS for three times, and these cells were fixed on coverslips by IF fixation solution for 15 minutes. Afterwards, the cells were permeabilised and stained simultaneously by 0.25% Triton X-100 with 1 in 250 diluted Phalloidin staining solution for 1 hour in room temperature with light protection. Subsequently, the cell-attached coverslip was washed with PBS and then was mounted with ProLong™ Gold /Diamond Antifade Mountant with DAPI. The cells on the coverslip were observed under a Leica SP5 confocal microscope after 24 hours incubation with mounting buffer. The fluorescent signals of the IF samples can be maintained for several months.

2.2.7 Cancer stem cells (CSCs) study

Different CSC identifying methods which point out various CSC characteristics have been developed over the decades. In breast cancer studies, quite a few methods have been recognised as CSC identifying methods. Three commonly used FACS methods including CD44 and CD24 cell surface antigens, aldehyde dehydrogenase (ALDH) activity, and cell efflux ability (Figure 2-8) and a functional CSC accumulation assay (mentioned in the section 2.2.1.2) (Kai et al., 2010) were utilised in this study. The CSC research in this project was mainly based on the results of FACS analysis unless otherwise specified.

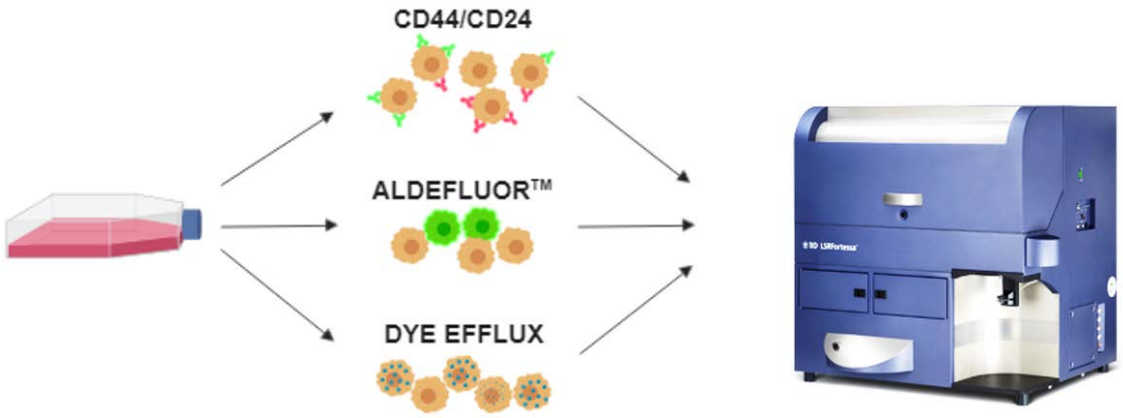


Figure 2-8 Three commonly used FACS-based BCSC identifying methods
The first commonly used method (on the top) is CD44/CD24 cell surface antigens staining method, cells are stained with CD44 and CD24 antibodies and then identified by flow cytometry, the cells with $CD44^{+}/CD24^{-}$ features are recognised as the *mesenchymal-like* CSCs. The second commonly used method (in the middle) is ALDH functional assay, cells with higher ALDH activities (green cells) can be identified by flow cytometry and are considered as the *epithelial-like* CSCs. The third commonly used method (on the bottom) is cell efflux ability assay, cells with great dye efflux capabilities can also be identified by flow cytometry and are considered as the side population cells.

2.2.7.1 CD44^{+}/CD24^{-} mesenchymal-like CSCs

The existence of CD44 and CD24 antigens on the cancer cell membrane has been employed as a valid CSC indicator for a long time. The exact function of CD24 has not been determined; whereas, CD44 has been linked with some unfavourable outcomes in cancers such as cancer metastasis and cancer proliferation (Hu et al., 2018). Evidence suggested that breast cancer cells with low CD24 and high CD44 expression can initiate tumour in immunodeficient mice (Velasco-Velazquez et al., 2012) which is a typical CSC characteristic.

For identifying the expression of CD44 and CD24 on the cell surface, the harvested cells were stained by an anti-CD44 antibody conjugated with PE fluorescence and anti-CD24 antibody conjugated with PE-Cy7™ or Alexa Fluor® 647 fluorescence depending on the fluorescent signals of the co-stained antibodies. The sample was incubated with the target antibodies in a 4°C fridge protecting from light for 30 minutes, followed by gentle wash steps. After a complete residue dye wash, the sample was passed through and detected by a flow cytometry BD Fortessa. The data was collected by the machine and was further analysed by Flowjo software version X.

2.2.7.2 High ALDH activity epithelial-like CSCs

Cancer cells with high ALDH activity have been recognised as CSCs and have been widely used in many studies. These cells also can initiate tumour in immunodeficient mice as CD44^{+}/CD24^{-} cells do (Ginestier et al., 2007). A commercially available functional CSC assay has been established by STEMCELL technologies named ALDEFLUOR® Kit for stem cell identification (Cat# 01700). The ALDH substrates expressed green fluorescence while the substrates were digested by ALDH. For identifying the cancer cells with high ALDH activity, the harvested cells were incubated with an appropriate amount of pre-mixed activated ALDEFLUOR® substrate in ALDEFLUOR® buffer for 30 to 40 minutes in a 37°C incubator to let ALDH digest the substrates. In the meantime, the same amount of cells were incubated with ALDEFLUOR® substrate within similar conditions to the test cells apart from the addition of N, N-diethylaminobenzaldehyde (DEAB) for blocking ALDH function. After incubation, the test sample and control cells were detected by a flow cytometry BD Fortessa. The data was collected by the machine and further being analysed by Flowjo software version X. The high ALDH activity cells were gated based on the ALDH inhibition control. Cells brighter than the ALDH inhibition control were considered as high ALDH activity cells.

2.2.7.3 Side population

Side population describes a small among of cancer cells having strong efflux ability associating with ABC transporters and maybe one of the causes of cancer drug resistance (Moserle et al., 2010). However, increasing evidence suggests that the side population might be unable to determine CSC and might only be responsible for a part of drug resistance (Britton et al., 2012). For determining side population, the harvested cells were incubated with Hoechst 33342 (Sigma-Aldrich, Cat# B2261) at a final concentration of 5 µg/ml for 90 minutes in a 37°C incubator for cells to be stained and/or efflux Hoechst dye. Afterwards, the sample was washed with DPBS buffer to remove the residue Hoechst dye followed by adding PI as an indicator of dead cells. Finally, the sample was passed through and detected by a flow cytometry BD Fortessa with dedicated settings that detect blue and red fluorescence of Hoechst dye. The data was collected and further being analysed by Flowjo software version X. Cells with a greater efflux ability can be identified from the major population, and these cells were considered as side population.

2.2.8 RNA sequence (RNA-Seq) analysis

The entire RNA-Seq process includes sample import and data export as depicted in Figure 2-9. Details of RNA-Seq preparation and preliminary data analysis was described below.

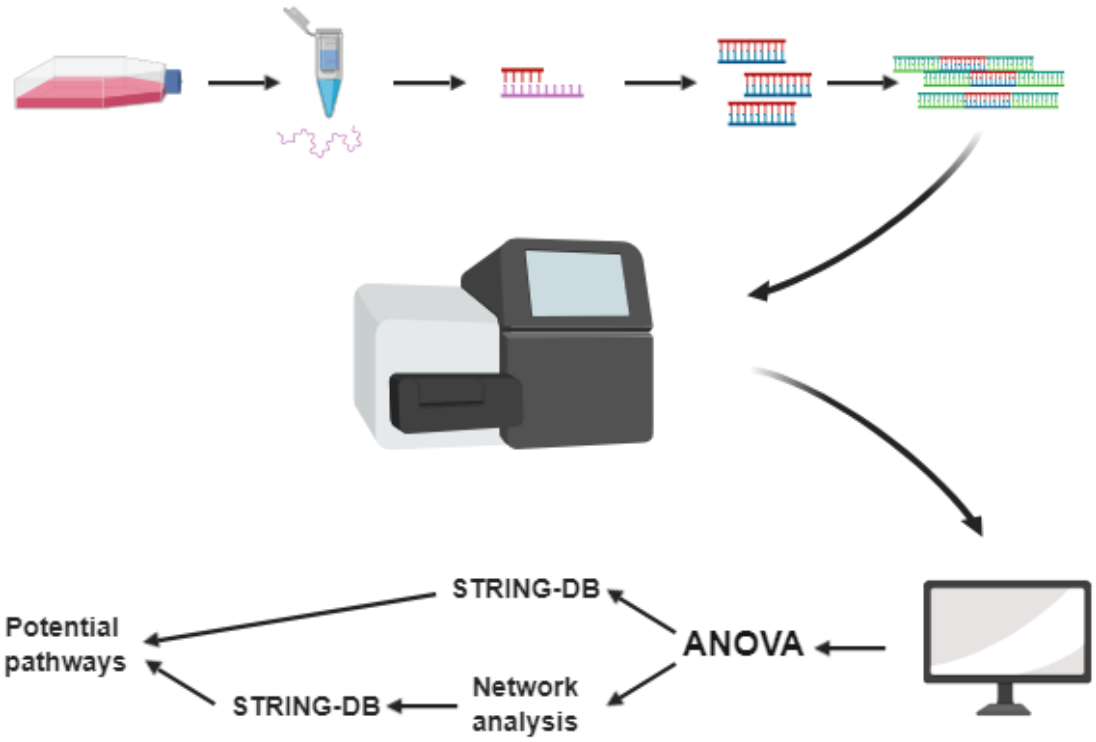


Figure 2-9 RNA-Seq process
High-quality total RNA extracted from samples are reverse transcribed to cDNA. RNA-Seq libraries are established based on these cDNA. A sequencer exports the RNA-Seq results after analysing the RNA-Seq libraries. The results are then further analysed by ANOVA analysis followed by STRING-DB analysis or network analysis in order to acquire potential molecular or pathways which may be involved in drug responses.

2.2.8.1 RNA-Seq preparation

RNA-Seq preparation can be separated into two major parts, one is sample preparation, and the other is sequencing and preliminary analysis.

The sample preparation of RNA-Seq requires high-quality total RNA as shown in Table 2-8 and Figure 2-10, the steps of total RNA isolation was described in section 2.2.7.1.

Table 2-8 The quality of total RNA for RNA sequencing

Sample	RIN value	28s/18s ratio
MCF10A DMSO	9.5	2.0
MCF10A EFV	9.0	2.2
MCF10A SPV	9.6	2.2
MCF10AT DMSO	8.7	2.3
MCF10AT EFV	8.5	2.1
MCF10AT SPV	8.9	2.2
MCF10CA1a DMSO	9.2	2.2
MCF10CA1a EFV	8.9	2.3
MCF10CA1a SPV	9.2	2.3
MDA-MB-231 DMSO	9.1	2.2
MDA-MB-231 EFV	10	2.6
MDA-MB-231 SPV	10	2.7
MCF7 DMSO	9.7	2.1
MCF7 EFV	9.9	2.1
MCF7 SPV	9.9	2.2

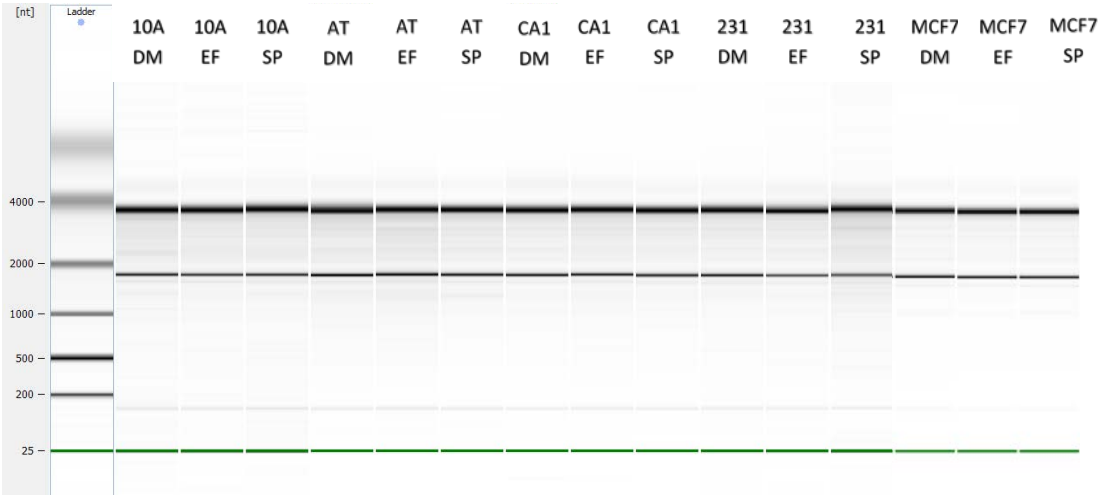


Figure 2-10 The quality of total RNA analysed by Aglient 2100 bioanalyzer for RNA sequencing
The top bands represent 28S RNA; the second bands represent 18S RNA; the third bands represent RNA fragments around 100 nt; the last green bands represent the bioanalyzer marker. Two clear and sharp bands (28S and 18S RNA) and a light band can be detected in a non-degraded pure total RNA sample.

After collecting high-quality total RNA, the RNA-Seq libraries were made using the Truseq stranded mRNA library kit. When the instrument presented all the gene reads for the entire human genome from different samples, the reads were mapped with a standard pipeline based on Ensembl database version 72 using Tophat-2.0.12-1 software. The raw mapped read counts were then converted into Fragments per Kilobase Million (FPKM) values by the cufflinks-2.2.1 programme in order to make further comparisons. Following mapping, FPKM values were imported into Partek Genomics Suite version 6.3 (Partek, St Louis, Missouri, USA) using the log(x + 1) transformation to convert all FPKM values to logged values for further analysis.

2.2.8.2 ANOVA analysis

After the RNA-Seq data was collected, many methods can be used to analyse the data. The first method we used was the analysis of variance (ANOVA) analysis. ANOVA analysis is the most widely used method to compare the variance to calculate whether the gene displays a significant difference. The experiment was analysed in Partek Genomics Suite using a 2-way ANOVA model based on the method of moments (Eisenhart, 1947). The cell type factor had 5 levels (MCF10A, MCF10AT, MCF10CA1α, MDA-MB-231, and MCF7) and the treatment factor also had 3 levels (EFV, SPV, and DMSO-control). The contrasts EFV-treated cells versus DMSO-untreated-control cells and the SPV-treated cells versus DMSO-untreated-control cells contrasts were calculated based on the principles of data analysis (Tamhane and Dunlop, 2000) by Dr Stephen Ohms.

Typically in an ANOVA analysis, a *p*-value less than 0.05 is considered as showing a significant difference. However, for the multigene analysis, the cut-off for a significant difference needs to be corrected, otherwise, the power could be too low to identify the truly significant genes. As such, the optimised *p*-value, *q*-value, was used for determining significant genes in this study. *Q*-value takes into consideration the false discovery rate (FDR) and has a higher chance of detecting significant differences (Li et al., 2007). A gene with a *q*-value less than the set cut-off was considered as a gene with significant change and can be further analysed by the STRING-DB pathway searching software version 10.5.

2.2.8.3 Network analysis

Network analysis for RNA-Seq results is based on interpreting the ANOVA results. The principle of this method is to gather the genes with a similar trend to a module. This module may contain some relevant genes since they have a similar expression pattern in different conditions. Thus, when analysing these genes by the STRING-DB pathway searching tool, there is a chance of finding genes which are correlated with each other.

The preliminary RNA-Seq data analysis including ANOVA analysis and network analysis was done by Dr Stephen Ohms. We sincerely appreciate his assistance and cooperation.

Statistical analysis

The data obtained from each experiment (immunofluorescence, western blot, RT-PCR, cell proliferation assay, cell cycle assay, CSC assays, and others) were expressed as the mean ± SD of three independent experiments with each experiment performed in triplicate. A student's t-test was used to estimate statistical significance using Prism software (GraphPad Software, San Diego, CA). $P < 0.05$ was accepted as statistically significant. If not, need to amend in each section where relevant

Chapter 3 The effects of antiretroviral drugs on triple-negative breast cancer cell lines

3.1 Introduction

Breast cancer is one of the most common and lethal diseases worldwide, however, some of its causes and regulatory pathways involved are beyond our understanding. On average, more than 50 people are diagnosed with breast cancer daily in Australia, most are women older than 50 years old (AIHW, 2017). In Australia, one in seven women has a chance of being diagnosed with breast cancer in her lifetime whereas, for men, the statistics are one in 675 (AIHW, 2017). Although the five-year survival rate of breast cancer has increased from 74% to 90.8% within the past few decades, it is still the fourth leading cause of cancer-related deaths in females in Australia (AIHW, 2017). The incidence and mortality of breast cancer reflect on our lack of knowledge and inadequate control of the disease. It is therefore very important that we further our understanding of the mechanisms involved in breast cancer initiation, growth, and progression.

Triple-negative breast cancer (TNBC) is a subclass of breast cancer with poor prognostic outcomes. About 15% of breast cancer patients are TNBC patients. TNBC lacks the oestrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2) (Reddy, 2011). Breast cancer patients differ in their prognostic outcomes because they respond differently to the drugs. Hormone therapies can only treat breast cancers with hormone receptors, and HER2 inhibitors are used for treating breast cancers with HER2 overexpression (Jamdade et al., 2015). Unlike breast cancers with well-known receptors, TNBCs are usually treated with systemic chemotherapy and radiotherapy (Collignon et al., 2016) because there are very few targeted therapeutic strategies available to them (Kalimutho et al., 2015).

One of the potential therapeutic targets for breast cancer treatment is the endogenous transposon, long interspersed nucleotide element 1 (LINE-1). In the mid-1990s, the incidence of *acquired immune deficiency syndrome (*AIDS)-related cancers was greatly reduced in human immunodeficiency virus (HIV)-infected patients owing to the introduction of antiretroviral therapy (Simard and Engels, 2010). Some scientists believed the decreasing incidence of AIDS-related cancers was not only due to the successful inhibition of HIV reverse transcriptase (RT) but also to the inhibitory effect of antiretroviral drugs on the autonomous LINE-1 RT activity

in tumour cells (Jones et al., 2008). LINE-1 is almost silent in differentiated tissues via several cell self-defence mechanisms because of its mutagenicity to normal differentiated cells; however, it is highly expressed in tumour cells (Sciamanna et al., 2016). LINE-1 expression has been identified in various cancers, including most breast carcinoma (97% LINE-1 positive cases in a screening study), as identified by screening studies of clinical samples (Rodić and Burns, 2013, Rodić et al., 2014). Furthermore, inhibition of LINE-1 has been shown to eliminate cancer cells in prostate cancer, melanoma and some non-TNBCs. Therefore the role of LINE-1 in TNBCs have been investigated in this study in order to improve targeted therapeutic strategies for TNBCs.

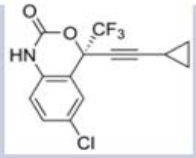
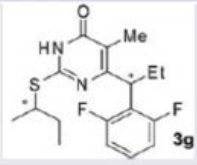
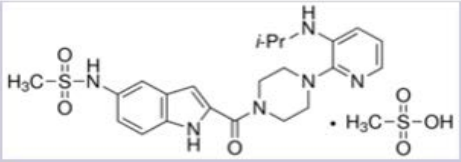
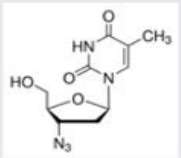
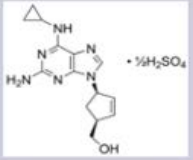
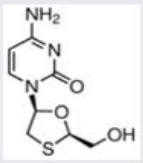
The focus of this chapter was to test antiretroviral drugs for their inhibitory effects on TNBCs. Whether LINE-1 plays a critical role in any observed inhibitory effects was investigated as was the mechanism involved. These experiments give a brief insight into the effectiveness of repurposing antiretroviral drugs as anticancer drugs in TNBCs.

3.2 Results

3.2.1 Antiretroviral drugs effective against breast cancer cells

The anticancer effects of certain antiretroviral drugs have previously been confirmed in epithelial cancer cell lines such as prostate, colon, lung, pancreatic cancers, and leukemia (Rodić and Burns, 2013). Some studies suggest that nucleoside reverse transcriptase inhibitors (NRTIs), or nucleoside analogs, may have more potential for eliminating cancer cells than non-nucleoside reverse transcriptase inhibitors (NNRTIs). Other studies, however, show NNRTIs to have significant anticancer effects (Jones et al., 2008). These results suggest that drug responses may vary depending on the cell type.

Table 3-1 List of the tested antiretroviral drugs and their structures

Drug	Abb.	Type	Structure
Efavirenz	EFV	NNRTI	
SPV122	SPV	NNRTI	
Delavirdine	DEL	NNRTI	
Zidovudine	AZT	NRTI	
Abacavir sulfate	ABC	NRTI	
Lamivudine	3TC	NRTI	

(NNRTI: non-nucleoside reverse transcriptase inhibitor; NRTI: nucleoside reverse transcriptase inhibitor)

In order to verify the anticancer effects of antiretroviral drugs on TNBCs, some antiretroviral drugs were selected and tested on a variety of human TNBC cell lines, using a cell viability assay. Some commonly used antiretroviral drugs including Efavirenz (EFV, NNRTI), Delavirdine (DEL, NNRTI), Zidovudine (AZT, NRTI), Abacavir sulfate (ABC, NRTI), Lamivudine (3TC, NRTI), and a newly synthesised drug SPV122 (SPV, NNRTI) (Table 3-1) were chosen for testing on MCF10AT, MCF10CA1α, and MDA-MB-231 TNBC cell lines. MCF10A cell line was a non-cancerous control, whereas the T47D cell line was selected as a luminal-type breast cancer control because T47D is highly sensitive to many antiretroviral drugs. The cytotoxic effects of the drugs were investigated on the chosen cell lines, using the XTT cell viability assay.

To investigate breast cancer cell viability under various treating conditions, different concentrations of the drugs EFV, SPV, DEL, AZT, ABC, and 3TC (0 µM, 10 µM, 20 µM, 30 µM, 40 µM, and 50 µM) were individually incubated with MCF10A, MCF10AT, MCF10CA1α, MDA-MB-231, and T47D cells. The cells were treated with the drugs for four days because the greatest variations in cell number were observed at the 96-hour time point of drug treatment (in appendix Figure S3-1). The cell viability results (Figure 3-1) indicated that EFV and SPV were the most effective anticancer drugs on all the breast cancer cell lines tested; DEL and AZT were less toxic than EFV and SPV; whereas 3TC and ABC under 50 µM were non-toxic (data are not shown) against all breast cancer cell lines tested. In contrast and as predicted, MCF10A, a non-cancerous cell line, showed no response to any of these drugs at concentrations less than 50 µM.

The half-maximal effective concentration (EC50) values of the drugs were calculated by Prism software (Table 3-2). After converting drug concentrations to logarithmic doses, the ‘log (inhibitor) versus normalized response – variable slope’ function in Prism was utilized to calculate EC50 values of the drugs by using the equation: ‘Y=100/(1+10^((LogEC50-X)*HillSlope)))’. The EC50 values of EFV in MCF10AT, MCF10CA1α, MDA-MB-231, and T47D were 23.32 µM, 21.84 µM, 27.89 µM, 9.19 µM, respectively; the EC50 values of SPV in MCF10AT, MCF10CA1α, MDA-MB-231, and T47D were 23.41 µM, 12.47 µM, 28.13 µM, and 6.57 µM, respectively. As expected, SPV was a slightly more effective anticancer drug than EFV in most of the cell lines except MCF10AT. This observation agreed with previous findings

in skin malignant melanoma carried out by the manufacturer of SPV, who designed it to precisely target LINE-1 RT, unlike EFV repurposed from an HIV inhibitor which may possibly lose its target on LINE-1 (Sbardella et al., 2011).

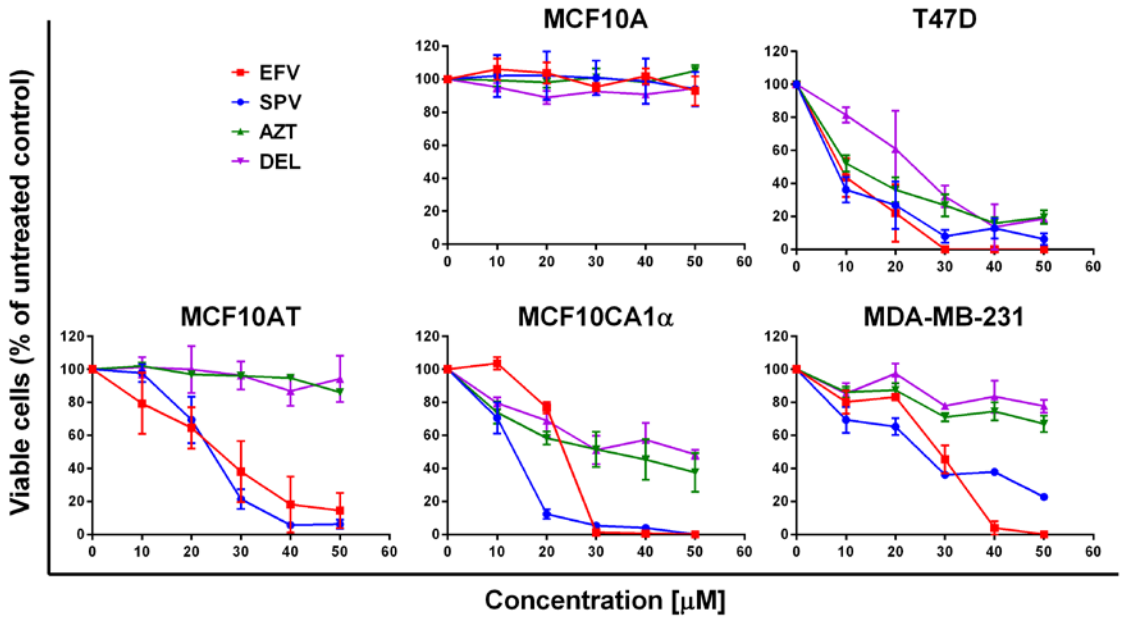


Figure 3–1 Cell viabilities at different antiretroviral drug concentrations
MCF10A gave no response to any of the tested drugs. EFV (red), SPV (blue), AZT (green), and DEL (purple) were toxic to T47D cells. TNBC cell lines – MCF10AT, MCF10 CA1α, and MDA-MB-231 cell lines had fewer responses to AZT and DEL compared with T47D cell line; whereas, EFV and SPV were more toxic to TNBC cell lines compared with other tested drugs. Among the six drugs tested, EFV and SPV were highly toxic for all tested breast cancer cell lines. Cell viability was determined at day 4 using a cell viability assay. Error bars: ±SD, n = 3.

Table 3-2 The EC50 values of EFV and SPV in each cell line

Drug	type	MCF10AT	MCF10CA1α	MDA-MB-231	T47D
Efavirenz (EFV)	NNRTI	23.32 μM	21.84 μM	27.89 μM	9.19 μM
SPV122 (SPV)	NNRTI	23.41 μM	12.47 μM	28.13 μM	6.57 μM

In summary, the EC50 values indicated that EFV and SPV are effective anticancer drugs for treating TNBC cell lines compared with other tested antiretroviral drugs. T47D luminal-type breast cancer cell line was sensitive to EFV, SPV as well as to AZT and DEL, and had lower EC50 values for all these drugs than did TNBCs. This illustrated that TNBCs could be more resistant to these antiretroviral drugs than luminal-type T47D cancer cells, yet were still affected by the drugs at relatively higher concentrations. The following experiments were based on the EC50 values for EFV and SPV for each cell line.

3.2.2 EFV inhibits LINE-1 expression

After determining EFV and SPV to be the most effective compounds against the selected TNBC cell lines, we questioned the role of LINE-1 in this process. Western blot data compared LINE-1 protein ORF1 and ORF2 expression in DMSO-untreated-control and EFV-treated MCF10AT, MCF10CA1α, MDA-MB-231, and T47D cancer cells as well as in non-cancerous MCF10A control cells (Figure 3-2). These data indicated that very limited LINE-1 protein expression was detected in the MCF10A cell line (Figure 3-2A); whereas all breast cancer cell lines expressed LINE-1 proteins (Figure 3-2A, 3-2B). Also, the data showed that less LINE-1 proteins were detected after EFV treatment compared with untreated-controls (Figure 3-2A, 3-2B). LINE-1 protein expression in the EFV-treated cancer cells was reduced to about 60% of that in untreated cancer cells (Figure 3-2D). In addition, after treating these breast cancer cells with EFV, ORF2, which is the protein containing reverse transcriptase region (the potential target of antiretroviral drugs), and ORF1 were suppressed in all of the cancer cell lines tested, suggesting an overall LINE-1 knockdown in EFV-treated breast cancer cells. Moreover, T47D western blot data indicated that the longer (six days) the cancer cells were exposed to EFV, the less LINE-1 proteins were expressed (compared with four days of EFV treatment) (Figure 3-2C).

Consistent with the western blot data, the qRT-PCR data suggested that the expression of both LINE-1 *orf1* RNA and *orf2* RNA (including endonuclease (EN) and reverse transcriptase (RT) regions) was reduced after EFV treatment in MCF10A, MCF10CA1α, MDA-MB-231, and T47D cell lines (Figure 3-3). These experiments implied that EFV might affect the expression and function of LINE-1 protein for generating novel LINE-1 RNA in the cancer cell by a positive feedback loop. This indicated a possible link between EFV treatment-induced cytotoxin and LINE-1 inhibition in breast cancer cells, further supporting our view that EFV could be a valid LINE-1 inhibitor in these breast cancer cell lines.

LINE-1 insertion in human genome relies on L1RNP which is assembled with LINE-1 RNA, LINE-1 ORF1 protein, and LINE-1 ORF2 protein (Goodier et al., 2013). Once LINE-1 RT is blocked by antiretroviral drugs, LINE-1 reverse transcription can be terminated leading to less new LINE-1 being present in the cancer genome. Additionally, dissociated LINE-1 RNAs can be degraded (Kemp and Longworth, 2015), and therefore reduced levels of LINE-1 protein and RNA is detected in these antiretroviral drug-treated cells.

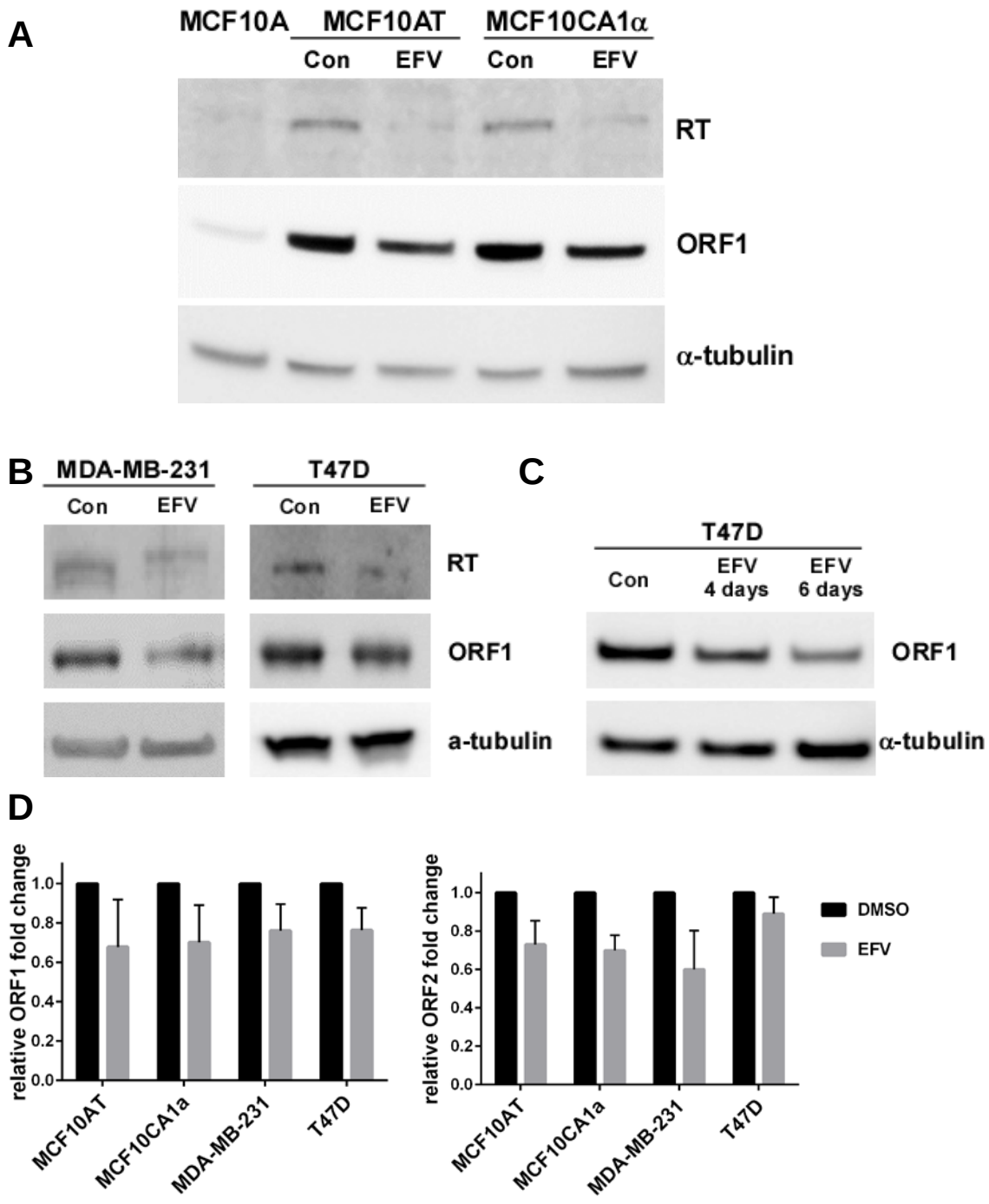


Figure 3–2 LINE-1 protein expression in different cell lines

Western blot data showing LINE-1 expression in breast cancer cell lines. (A) Almost no LINE-1 expression in non-cancer MCF10A cell line. LINE-1 proteins (ORF1 and ORF2 RT) were detected in MCF10AT and MCF10CA1α, but the expression levels were reduced by EFV treatment. (B) LINE-1 proteins were also detected in MDA-MB-231 and T47D cell lines and were inhibited by EFV treatment. (C) Increasing the EFV treatment period further inhibited LINE-1 expression. (D) The relative ORF1 and ORF2 RT fold-changes reflected LINE-1 inhibition by EFV in cancer cell lines. Error bars: ±SD, n = 3.

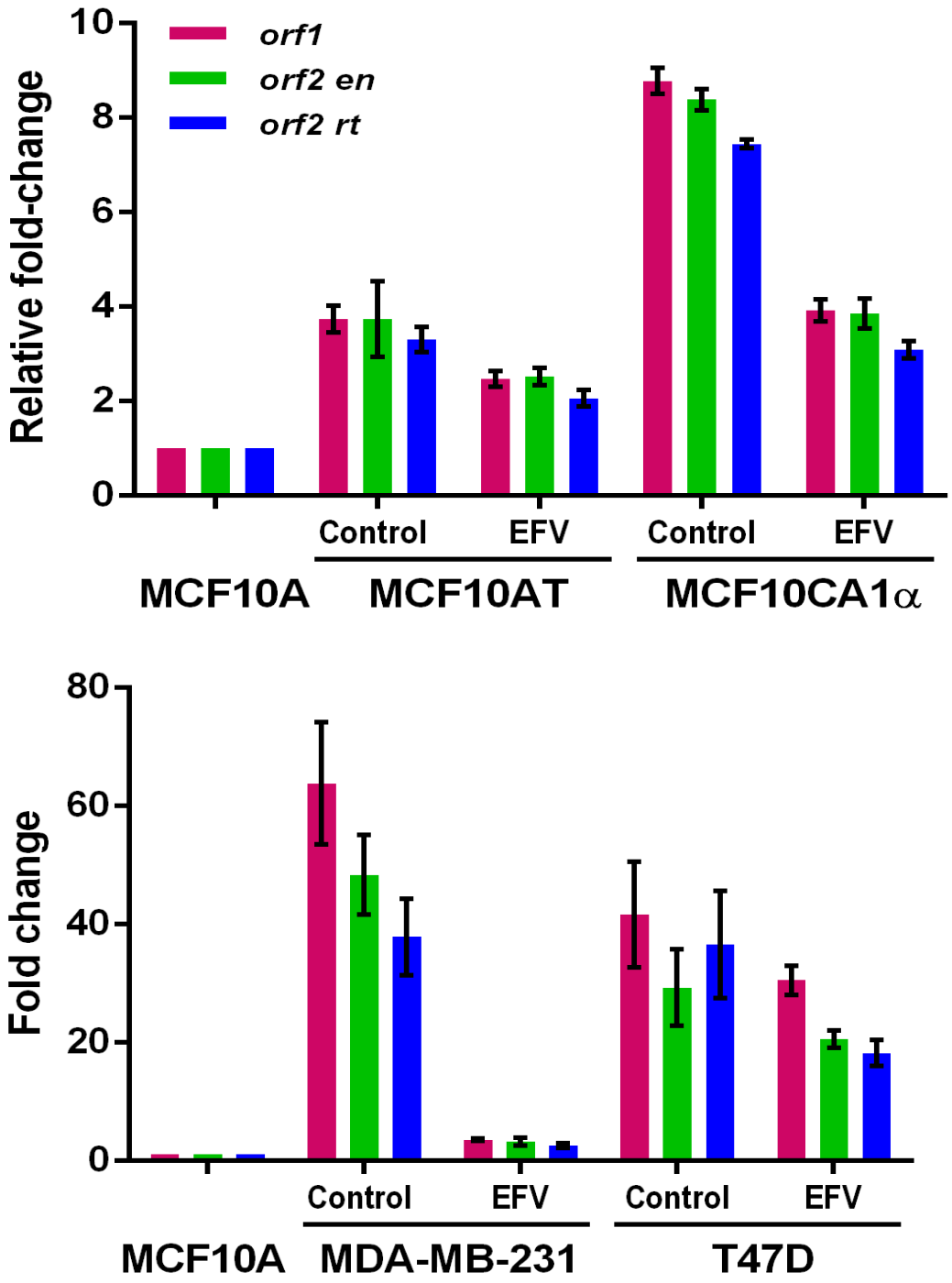


Figure 3–3 LINE-1 RNA expression in different cell lines
The qRT-PCR data showed that LINE-1 *orf-1* and *orf-2* (including endonuclease and reverse transcriptase regions) were expressed in breast cancer cell lines including MCF10AT, MCF10CA1α, MDA-MB-231 and T47D. EFV-treated breast cancer cells detected lower levels of LINE-1 RNA *orf-1* and *orf-2* compared with DMSO-controls. Error bars: ±SD, n = 4.

3.2.3 Antiretroviral drugs inhibit LINE-1 reverse transcriptase activities

3.2.3.1 LINE-1 overexpression

To further confirm the functional link between EFV and SPV anticancer effects and LINE-1 inhibition, a LINE-1 RT functional activity assay was performed in a non-cancerous and low-LINE-1-expressed cell line MCF10A. pBS-L1PA1-CH-mneo (Figure 3-4A) (Wagstaff et al., 2011) is an optimized LINE-1 overexpressing plasmid which can be transfected into a mammalian cell to express a target protein, in this case, optimized LINE-1. Its amino acid sequence is 74% consistent with the original LINE-1 sequence ensuring its escape from the cells' self-defence mechanisms while maintaining most of the LINE-1 characteristics. It has been shown to be capable of performing reverse transcription in cells (Wagstaff et al., 2011). The complete map of pBS-LAPA1-CH-mneo plasmid DNA was illustrated in section 2.1.7.

In this experiment, only the successfully pBS-L1PA1-CH-mneo transfected MCF10A cells survived in culture medium containing 200 μg/ ml Geneticin™ (Figure 3-4B1), which is usually harmful to normal (untransfected) cells (Figure 3-4B7, 3-4B8, 3-4B9). LINE-1 reverse transcriptase activities produce Geneticin™ resistance for the host cells. This Geneticin™ resistance was suppressed by antiretroviral drug treatment thus less antiretroviral drug-treated transfected cells survived (Figure 3-4B2, 3-4B3) compared with untreated-transfected cells (Figure 3-4B1). To quantify the numbers of viable cells in each condition, the cells were stained with crystal violet solution and de-stained with methanol. The de-stained crystal violet-methanol solution was collected and the optical densities of each flask measured at 570nm. A significant decrease was shown in EFV or SPV treated-transfected cells compared with their control (Figure 3-4C). The LINE-1 RT functional assay confirmed the effectiveness of EFV and SPV at inhibiting LINE-1 reverse transcriptase.

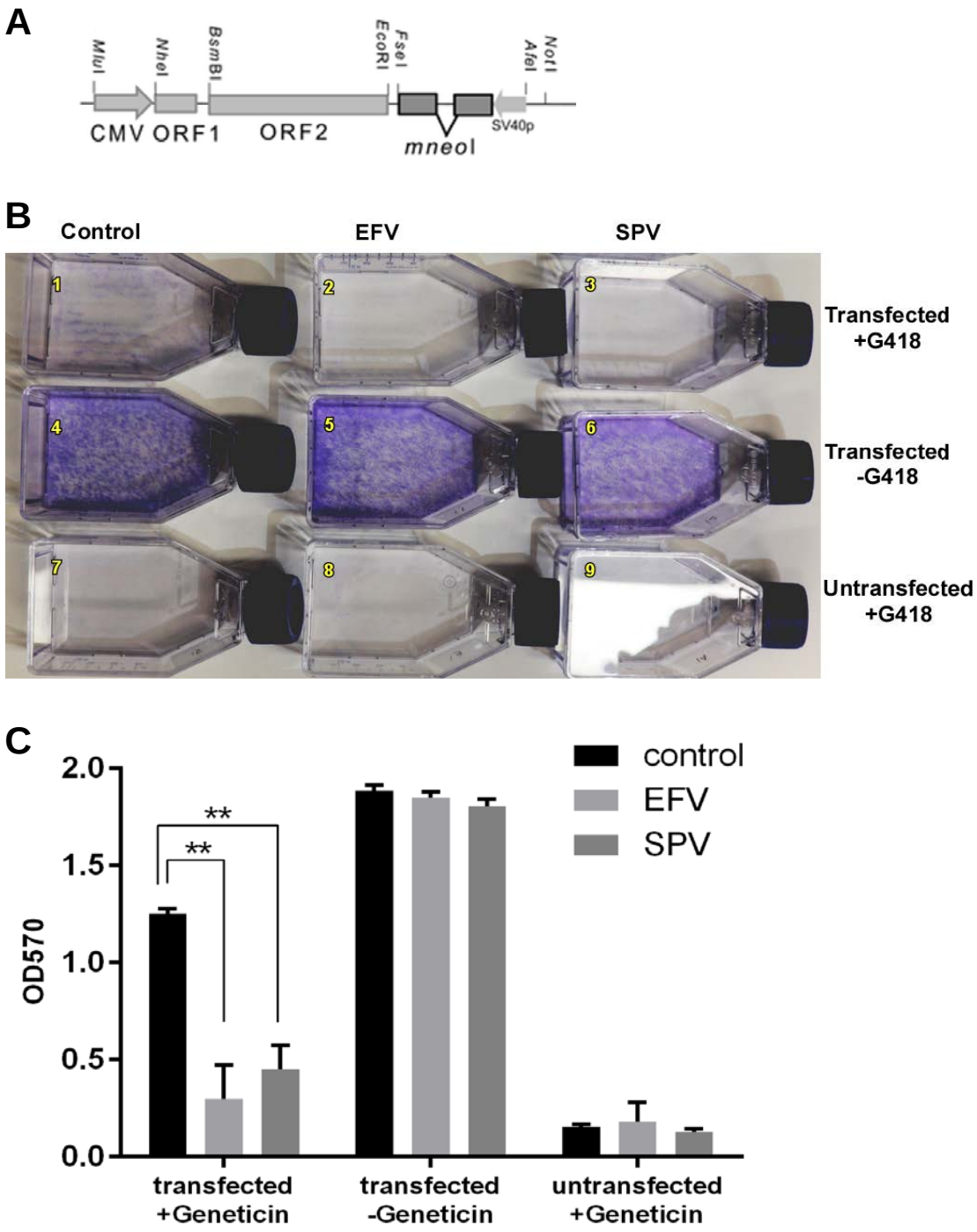


Figure 3–4 LINE-1 reverse transcription simulation assay in MCF10A cells
The pBS-L1PA1-CH-mneo plasmid (A) transfected cells have Geneticin resistance enabling their survival under Geneticin selection (200 µg/ ml of Geneticin) (B1); whereas, the untransfected control cells were all killed by 200 µg/ ml of Geneticin (B7, B8, B9). The cells' resistance was reduced by treating them with EFV or SPV (B2, B3). (B4, B5, B6 was no Geneticin selection control). Error bars: ±SD, n = 3. (Transfected cells with Geneticin selection, EFV-treated: p = 0.0082, SPV-treated: p = 0.0052. ** p-value < 0.01, two-tailed paired Student's t-test)

3.2.3.2 LINE-1 RNA inhibition

The experiments described in the previous section revealed a relationship between LINE-1 inhibition and the anticancer effects of EFV and SPV; however, the results did not confirm a direct link between them. Therefore, an RNA inhibition experiment was designed to compare the similarity of the phenotypes resulting from direct RNA inhibition and inhibition of LINE-1 by antiretroviral drugs.

In order to generate LINE-1 RNA inhibition cell lines, the pUTR plasmid (Figure 3-5B) encoding short hairpin RNA targeted towards the LINE-1 element promoter, its negative control, and the pSM2 plasmid (Figure 3-5A), which is an empty vector, were transfected into MCF10AT and MCF10CA1α cells using the Nucleofector® I machine. Since the pUTR and the pSM2 plasmids contain Puromycin resistance genes, those transfected cells had been culturing in the complete culture medium with 200 µg/ml Puromycin to allow selection of successfully transfected cells. More details about this shRNA inhibition system were described in Chapter 2.

To examine the consequence of shRNA inhibition, western blot was utilized to confirm the shRNA efficiency. Two critical cancer cell morphological indicators, fibronectin and E-cadherin, were stained in transfected cells to allow inspection of the morphological variations between LINE-1 inhibited TNBCs and their uninhibited controls.

Surprisingly, after ten days of Puromycin selection, the pUTR-transfected cells expressed marginally higher E-cadherin (an epithelial marker) and fibronectin (a mesenchymal marker) than the pSM2-control, in accordance with the immunofluorescence data (Figure 3-5D). However, no obvious morphological differences were observed between the pUTR-transfected cells and the pSM2-control cells in both MCF10AT and MCF10CA1α cell lines. Additionally, the western blotting result suggested that LINE-1 RNA inhibition was extremely low in MCF10CA1α cells: a very limited decrease of LINE-1 RT but not LINE-1 ORF1 was detected in the pUTR-transfected cells (Figure 3-5C). Thus, the comparison between direct RNA inhibition and inhibition of LINE-1 by antiretroviral drugs was unsuccessful. Possible explanations of these results are discussed in Chapter 6.

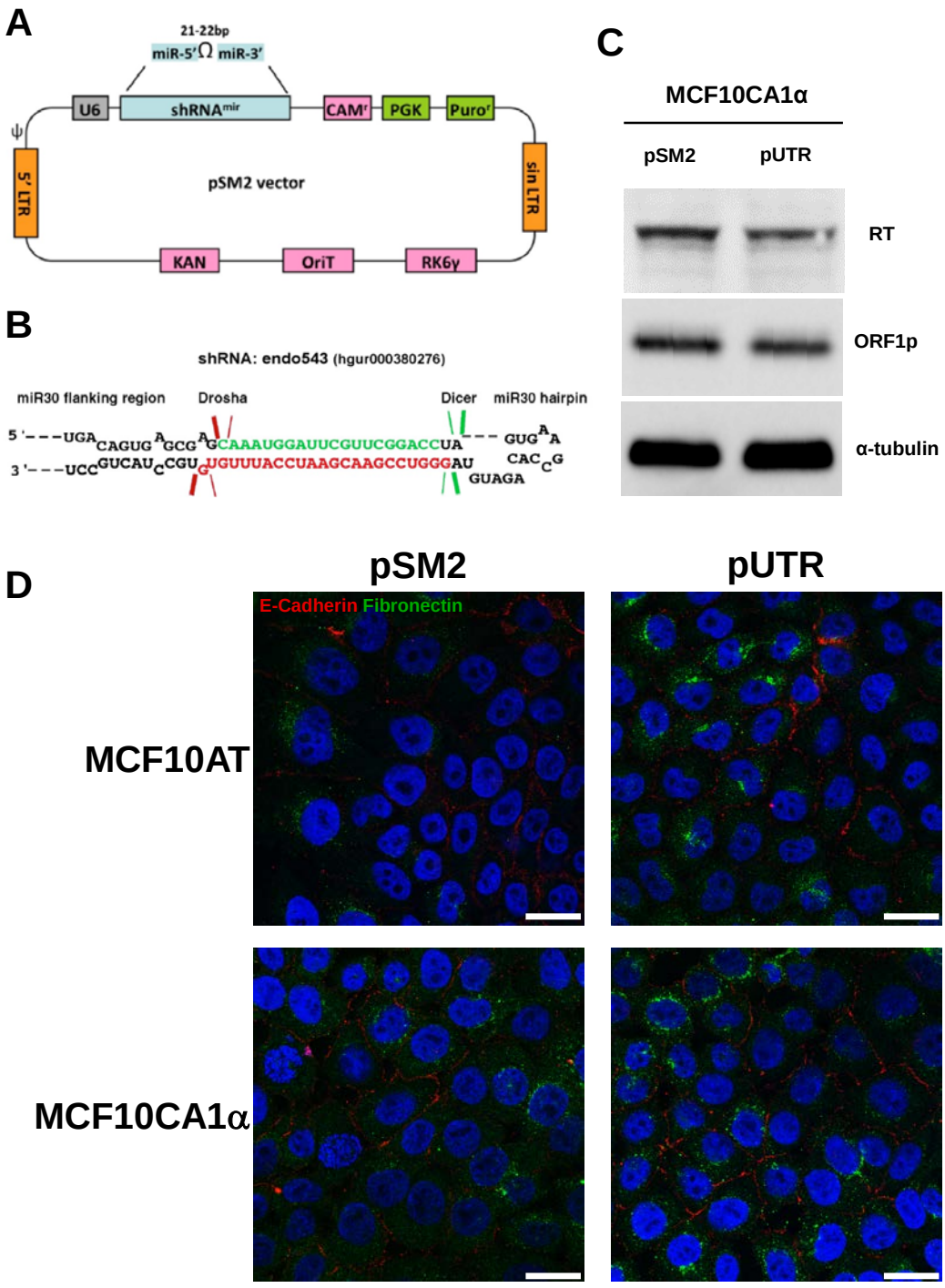


Figure 3–5 Morphological phenotypes in LINE-1 RNA silencing cells
(A) The pSM2 plasmid contains a puromycin resistant region and an shRNA insertion region. (B) The shRNA sequence can target the LINE-1 promoter and silence LINE-1. (C) The LINE-1 protein expression levels in the pUTR-transfected cells and the pSM2-transfected cells. α-tubulin is internal control. (D) The pUTR and the pSM2 plasmids were transfected into MCF10AT and MCF10CA1α cells. No obvious difference between the pUTR plasmid transfected cells and the pSM2 plasmid transfected cells was observed. Red dye stained E-cadherin; green dye stained fibronectin; blue dye stained cell nucleus. (Scale bar: 25 μm)

3.2.4 The influences of antiretroviral drugs in cancer cell growth.

After confirming the effect of antiretroviral drugs on the reduction of TNBCs, we investigated the mechanisms involved. Cancer is a severe disease that is strongly associated with the dysregulation of somatic cell proliferation (Evan and Vousden, 2001). Cell proliferation is defined as the process of increasing cell numbers and takes into account cell division, cell death, and other mechanisms changing cell numbers (Iván, n.d.). An uncontrollable increase in cell numbers may potentially generate a tumour. Therefore, reducing cancer cell proliferation might be an effective strategy for eliminating cancers. In this study, we have shown that antiretroviral drugs are capable of affecting cancer cell proliferation in XTT cell viability assay, and so we investigated whether cell division, cell cycle, and cell death were involved.

The XTT cell viability assay was first performed to observe the changes in the cell proliferation rate between untreated-control cells, EFV-treated cells, and SPV-treated cells at multiple time points for each cell line. This experiment (in appendix Figure S3-2) demonstrated that EFV treatment and SPV treatment could appropriately reduce the cell proliferation rate to half that of the untreated-control in all breast cancer cell lines. However, the XTT cell viability assay can only measure the relative numbers of live cells and the results can be influenced by cell death, hence the actual cell proliferation rate is uncertain.

3.2.4.1 Cell division

Therefore, we used the CFSE (Carboxyfluorescein succinimidyl ester) cell division tracking kit (Thermo Fisher Scientific) to confirm the XTT cell proliferation results. The CFSE kit utilizes a fluorescent molecule, carboxyfluorescein, to observe cell division. Carboxyfluorescein succinimidyl ester can be assimilated to cells and generates covalent coupling with amine sources within them, thus continues to fluoresce in cells for a long period of time. Once cell division has occurred, the brightness of the fluorescence reduces by half owing to the fluorescent carboxyfluorescein being equally divided between the two daughter cells. Hence, the brighter cell can be determined as an older cell which has gone through fewer cell divisions and has a lower cell proliferation rate.

In this experiment, the brightness of the cells was negatively correlated with the cell proliferation rate of the cell lines (Figure 3-6). The darkest MCF10CA1α had the highest cell

proliferation rate, whereas the brightest T47D had the lowest cell proliferation rate. The proliferation rate of MCF10AT was slightly lower than MCF10CA1α but higher than MDA-MB-231. Untreated-control cells divided faster than EFV-treated cells in all cell lines since brighter fluorescence was detected in EFV-treated cells compared with their controls. This indicated that EFV affects cell proliferation in these breast cancer cell lines. To further understand the possible reasons for the anti-proliferative effect of EFV, cell cycle and cell apoptosis experiments were performed.

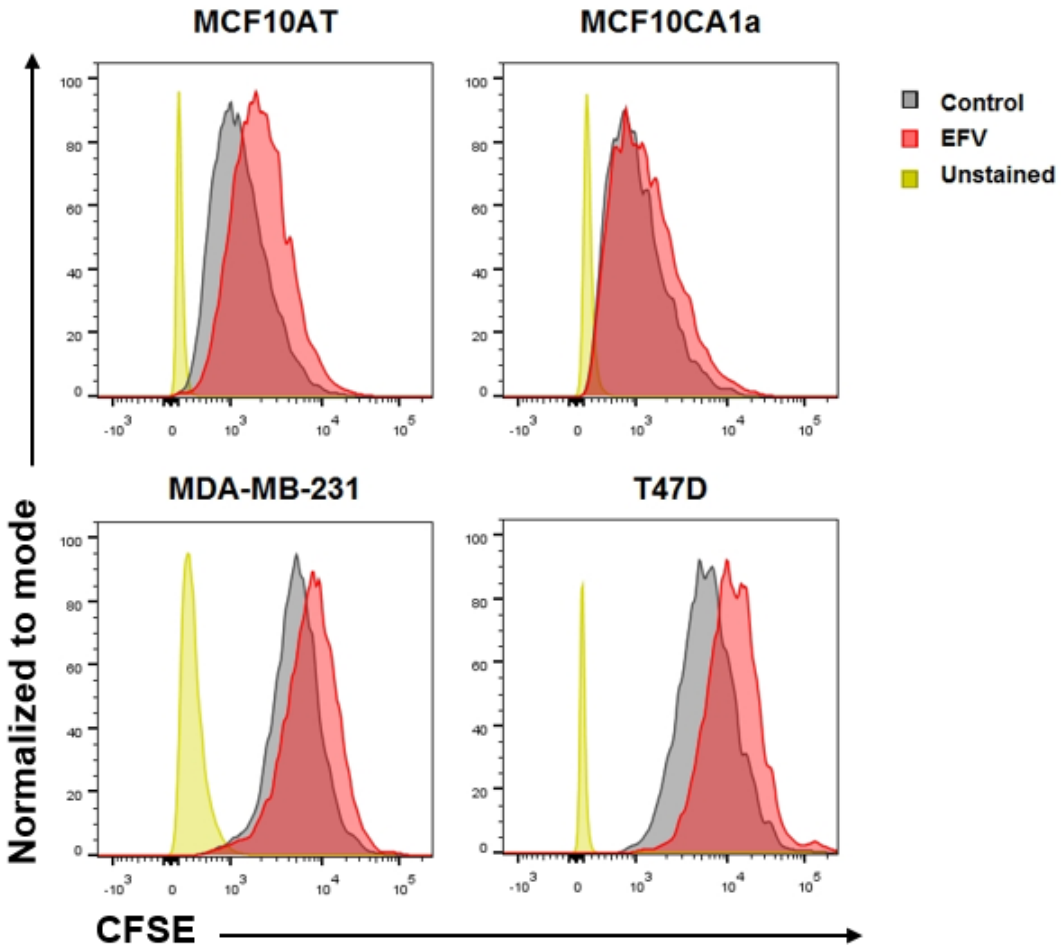


Figure 3–6 Cell division rates in EFV-treated cells and untreated-control cells
This experiment is to determine how fast cell division occurs in EFV-treated cells and untreated-control. The dark-yellow peak represents the unstained control, whereas the dark peak represents the untreated-control which had uptake CFSE before the culture period; the red peak represents the EFV-treated cells which also up took CFSE at the same time as the untreated-control cells. The CFSE is degraded through the passages, i.e., more the cell division show less fluorescence, which is closer to the dark-yellow peak in this FACS data.

3.2.4.2 Cell cycle

After observing the effect of EFV at reducing cell division with the CFSE assay, we investigated how the drugs influence the cell cycle. The samples were fixed by 80% ethanol on the fourth day of EFV treatment and then stained with Propidium iodide (PI), followed by fluorescence-activated cell sorting (FACS) analysis which can identify the brightness of the cells. Since PI can bind to DNA, the brighter cells contain more DNA. Cells in G1/G1 phase have 23 pairs of chromosomes; cells in S phase, which are in the synthesis stage have slightly more DNA than in G0/G1 phase; and cells in G2 phase, which are ready for cell division, have 46 pairs of chromosomes. Thus, PI staining can identify the different stages of the cell cycle.

Cell cycle data revealed some differences in Gap0 and Gap 1 (G0/G1), Synthesis (S) and Gap 2 (G2) phases between untreated- and EFV-treated cells in all cell lines (Figure 3-7). The results show the average of three independent experiments, the statistical results of Student's *t*-test were listed in the appendix Table S3-1. Comparisons between untreated- and SPV-treated cells was also noted, and the trend of cell cycle shifting in SPV-treated cells was similar to that in EFV-treated cells. Only a single successful experiment is shown due to the shortage of the SPV compound (in appendix Figure S3-3).

Not surprisingly, MCF10A cells showed very similar cell cycle patterns in antiretroviral drug-treated and untreated-control cells. Except for the MCF10A cell line, all the breast cancer cell lines had a lower proportion of S phase cells after drug treatment (MCF10AT: reduced from 11.42±2.56% to 6.02±3.63%; MCF10CA1α: reduced from 20.39±2.79% to 10.81±1.19%; MDA-MB-231: reduced from 21.52±3.99% to 18.81±5.74%; T47D: reduced from 19.46±3.46% to 15.70±0.17%), however, the changes varied in G0/G1 and G2 phase depending on the cell line. The EFV-treated MDA-MB-231 cells (increased from 23.53±5.02% to 29.28±6.01%) and T47D cells (increased from 21.70±2.17% to 25.80±4.32%) had higher proportions of G2 phase cells compared with their untreated-control cells; in contrast, the EFV-treated MCF10AT cells (reduced from 16.94±6.66% to 11.76±10.93%) and MCF10CA1α cells (reduced from 28.13±3.58% to 25.71±1.20%) had a lower proportion of G2 phase cells compared with their untreated-control cells. Altogether, these results indicated that DNA replication is reduced, and cell cycle retardation may occur during drug treatments.

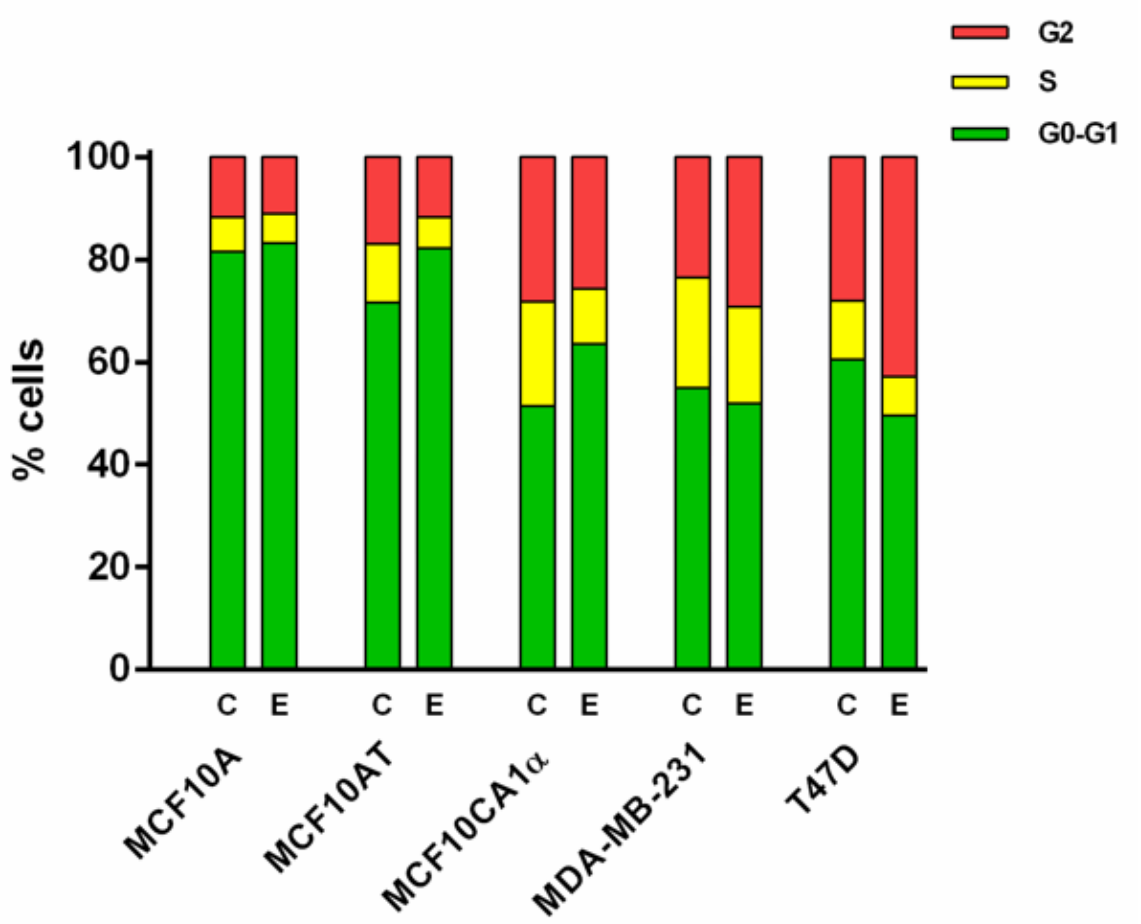


Figure 3–7 Cell cycle patterns of untreated-control and EFV-treated cells in MCF10A, MCF10AT, MCF10CA1α, MDA-MB-231, and T47D cell lines
The percentage of cells in different phases when exposed to different conditions was calculated and is presented in this figure. The green bar represents the percentage of cells in G0/G1 phase; the yellow bar represents the percentage of cells in S phase; the red bar represents the percentage of cells in G2 phase. C is the untreated-control cells; E is the EFV-treated cells. n = 3 for each cell line.

The proportion of subG1 phase cells (late stage of apoptosis) in the different treatment conditions was also investigated (in appendix Figure S3-4). They were raised in almost all EFV-treated cell lines except MCF10A. This suggests that apoptosis might be responsible for reducing cell numbers in antiretroviral drug-treated breast cancer cells. However, some papers recommend separating subG1 detection experiments from cell cycle experiments because the accumulation of apoptotic cells may affect the cell cycle results (Wlodkowic et al., 2011). Thus, the conclusions drawn from this experiment need to be reconsidered and we might need to perform further experiments to confirm the results.

3.2.4.3 Cell death

The Annexin V/ Propidium iodide (PI) apoptosis detection assay was then used to clarify the role of cell programmed death (apoptosis) and non-programmed death (necrosis) in antiretroviral drug-treated breast cancer cells. Annexin V is impermeable and can bind to phosphatidylserine (PS). PS can only bind to Annexin V when it translocates to the membrane surface i.e., the cell has undergone apoptosis or necrosis. The cell membrane is impermeable to PI, not PI impermeable to the cell membrane, and PI can only stain nucleotides if the cell membrane collapses. Therefore, cells with Annexin V positive and PI negative features are likely to be early apoptotic cells; whereas, cells with Annexin V positive and PI-positive features are probably late apoptotic and necrotic (Figure 2-7).

In this experiment, the data showed that after EFV treatment, the proportion of apoptotic and necrotic cells increased, compared with the untreated control cells (Figure 3-8). After excluding the majority of dead cells, the apoptotic and necrotic cells increased from 4.17% to 10.08% in MCF10AT; from 4.53% to 9.04% in MCF10CA1α; from 4.14% to 5.34% in MDA-MB-231, and from 3.58% to 9.86% in T47D. The data suggested that EFV treatment can induce cell apoptosis and necrosis in breast cancer cells.

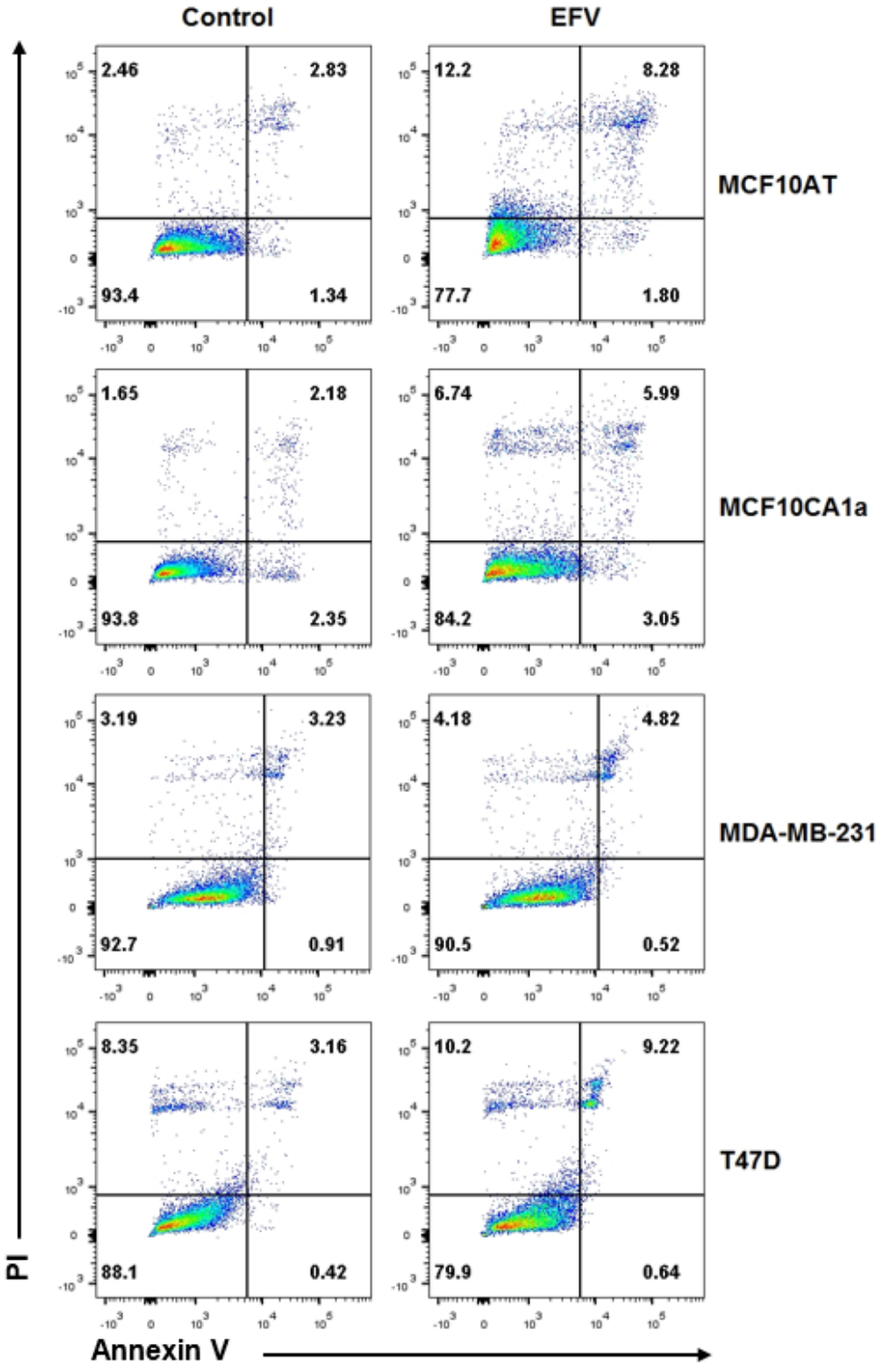


Figure 3–8 Cell apoptosis in untreated-control and EFV-treated breast cancer cells
This experiment detected the apoptotic cells and necrotic cells under different treating conditions by flow cytometry. Annexin V^+/PI^+ cells are late apoptotic and necrotic cells; Annexin V^+/PI^- cells are early apoptotic cells; Annexin V^-/PI^- cells are live cells; Annexin V^+/PI^+ cells are dead cells.

3.2.5 The effects of antiretroviral drugs on cell morphology

Cell morphology is an important indicator for distinguishing cancer cells from normal cells. Normal breast cell lines usually display differentiated epithelial phenotypes. In contrast, many breast cancer cell lines present undifferentiated and migration phenotypes associated with cancer invasion and metastasis (Bravo-Cordero et al., 2012). In addition, mesenchymal-like phenotypes can be observed which are strongly correlated with cancer malignancy (Zhang et al., 2017). The morphological phenotypes have been investigated with Phalloidin, a powerful F-actin probe, on untreated- and antiretroviral drug treated-cells in order to confirm whether the drugs can reduce the malignancy of TNBCs. The methods used for Phalloidin staining were described in Chapter 2.

As predicted, some morphological changes were observed in the drug treated-cells. In Figure 3-9, antiretroviral drug-treated breast cancer cells displayed cell projections due to elongated microtubules, representing cell differentiation (more images in Appendix Figure S3-5). The results showed that LINE-1 inhibition could potentially induce cellular differentiation, returning breast cancer cells to their original epithelial states. They also suggested that LINE-1 expression could be important for cancer cell mesenchymal to epithelial transition (MET). There were a few inconsistencies between cell lines. In MDA-MB-231 and T47D cell lines, cell differentiation was evident. There were a significant number of changes in the differentiation of cells observed between untreated- and EFV-treated cells of these cell lines (Figure 3-10). However, in MCF10AT and MCF10CA1α cell lines, some drug-treated cells were quiescent, some displayed cell scattering and an elongated shape, while others displayed lamellipodia, filopodia, and observable migratory behaviour (fan shape with clear direction) (Figure 3-11). These phenotypes could be associated with epithelial to mesenchymal transition (EMT) and be linked to unfavourable prognostic outcomes (Fedele et al., 2017). Thus, we were unable to confirm any enhancement of cell differentiation by antiretroviral drugs. This raised the following questions: Do antiretroviral drugs induce EMT? Would antiretroviral drug treatment lead to worse outcomes for patients?

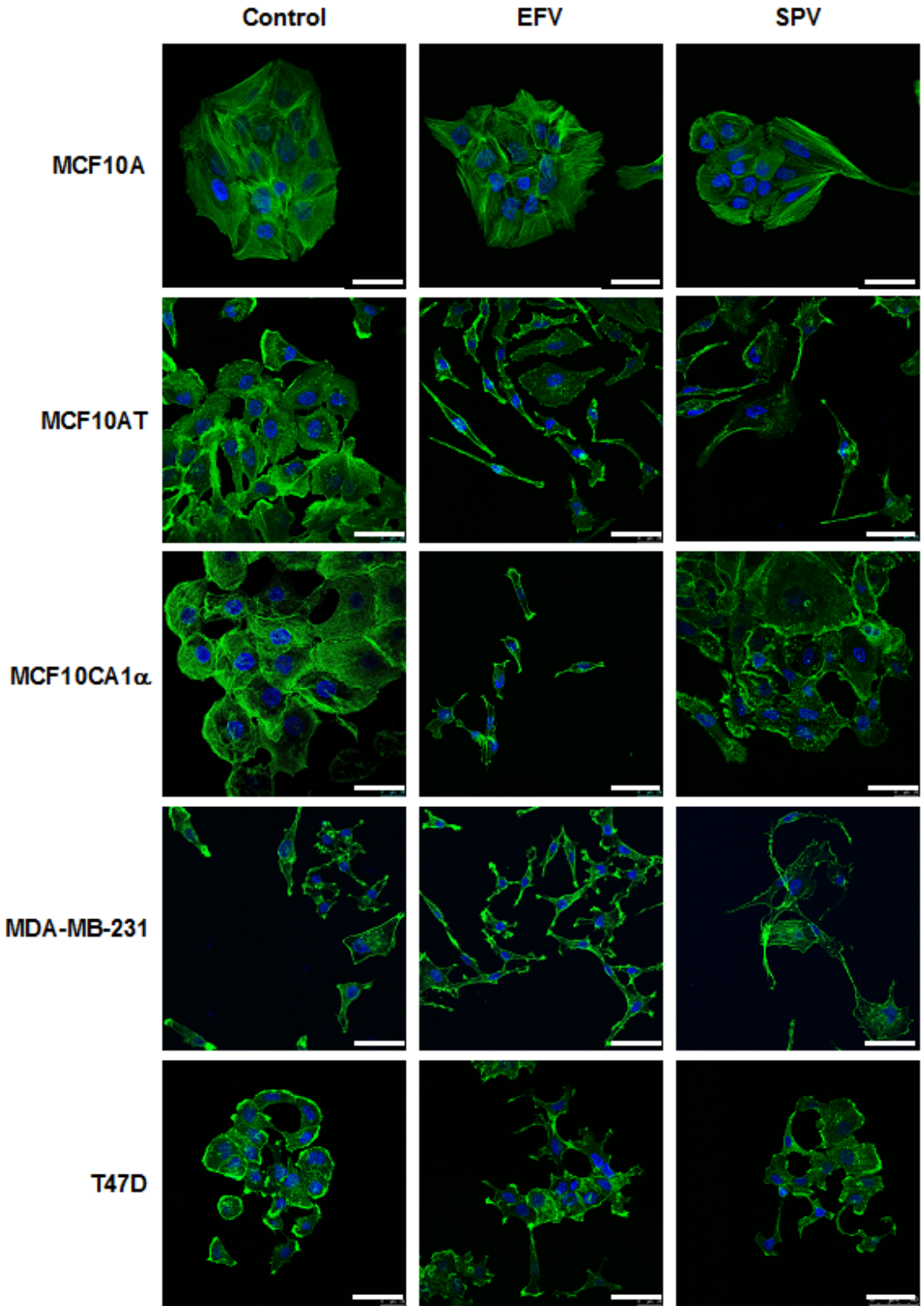


Figure 3–9 Cell morphologies in EFV- and SPV-treated and untreated cells
This figure compares the morphologies of antiretroviral drug-treated and untreated cells in MCF10AT, MCF10CA1α, MDA-MB-231, and T47D cell lines. F-actin (green, for detecting F-actin) and DAPI (blue, for detecting nucleus) were stained by Phalloidin. (Scale bar: 50 μm)

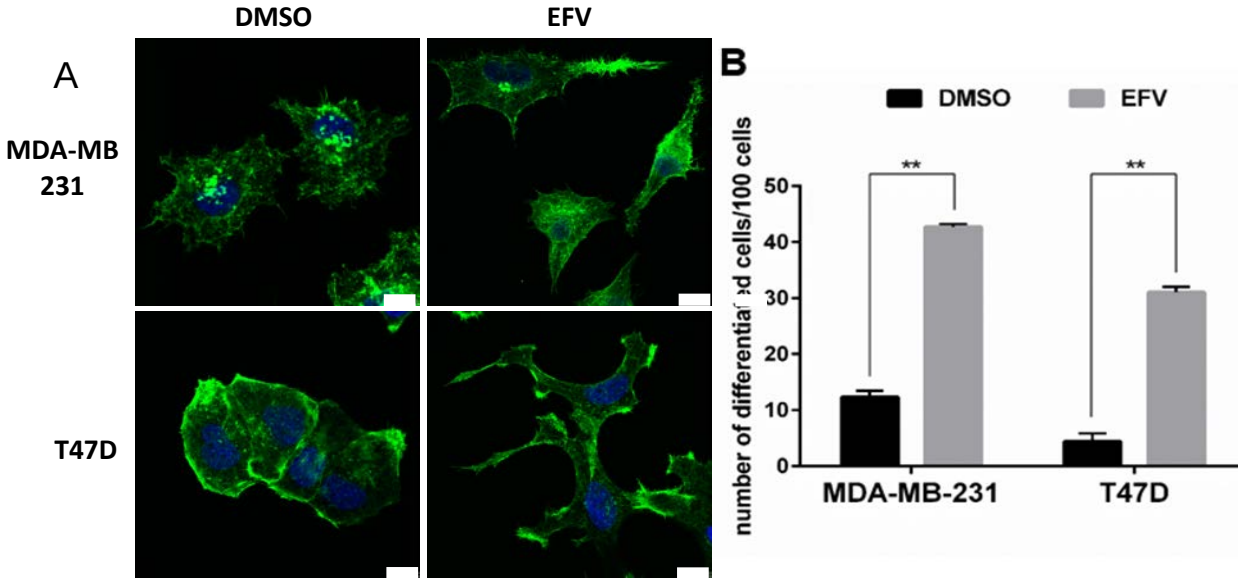


Figure 3–10 Differentiated cells in EFV-treated MDA-MB-231 and T47D cell lines
(A) immunofluorescence images for observing differentiated cells in MDA-MB-231 and T47D cell lines (Scale bar: 10 μm) (B) Numbers of differentiated cells in MDA-MB-231 and T47D cell lines. Error bars: ±SD, n = 3. (MDA-MB-231 p = 0.0008, T47D p = 0.0025. **p-value <0.01. Two-tailed paired Student's t-test.)

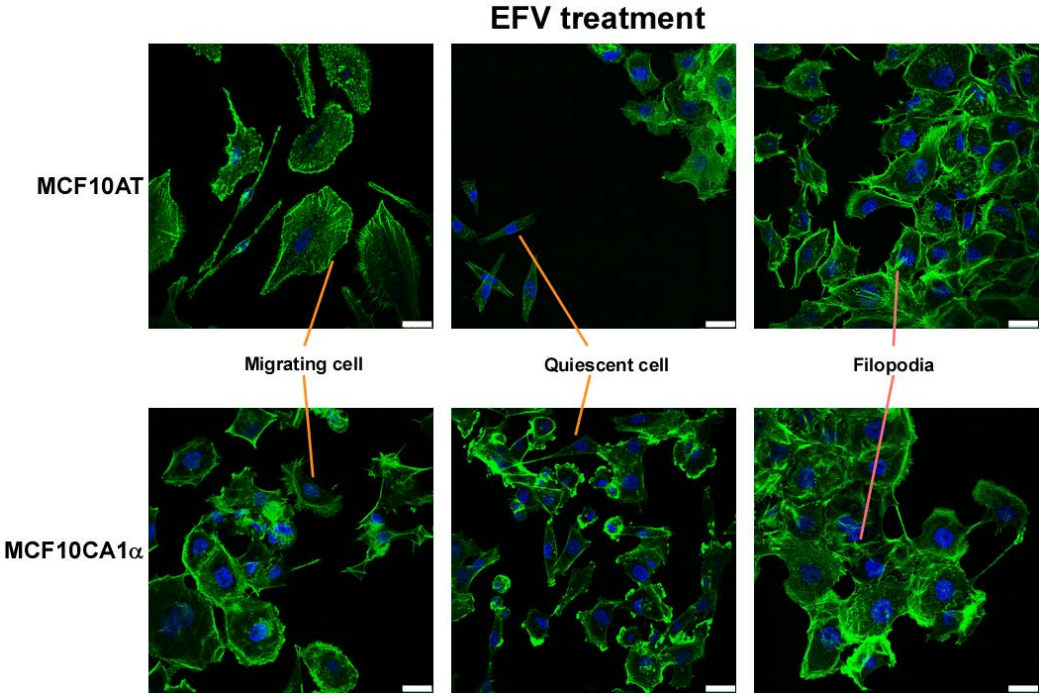


Figure 3–11 Different morphologies of EFV-treated MCF10AT and MCF10CA1α cells
After antiretroviral drug treatment, different cell morphologies can be observed in MCF10AT and MCF10CA1α cells. Migrating cells show clear filopodia and clear direction. Quiescent cells are usually be considered as potential CSCs and may cause cancer metastasis and cancer relapse. (Scale bar: 25 μm)

3.2.6 EFV treatment affects the expression of epithelial and mesenchymal markers in MCF10AT and MCF10CA1α cells

To help understand the reasons why MCF10AT and MCF10CA1α cells present multiple morphological changes after antiretroviral drug treatment, we used immunofluorescence and western blot to examine some well-studied cell morphological markers. E-cadherin is the most commonly used epithelial phenotype marker. It is considered to play a role in tumour suppression and has a critical role in the maintenance of the epithelial phenotype (Pećina-Slaus, 2003). Fibronectin and Snail Family Transcriptional Repressor 2 (SLUG or SNIL2) are confirmed mesenchymal phenotype markers, with relevance to tumour invasion and metastasis (Wang and Hielscher, 2017). Cell division control protein 42 homolog (CDC42) is a cell cycle regulator (Melendez et al., 2011). Lack of CD24 has been recognised as a marker for cancer stem cells (CSC), and DEAD-box protein 5 (DDX5) dysregulation is associated with cancer development (Wang et al., 2012).

According to the results shown in Figure 3-12, the immunofluorescence experiments suggested that more EFV-treated MCF10AT cells expressed more fibronectin and SLUG than untreated control cells. These proteins are the mesenchymal markers for mammalian cells. Less CD24 protein expression and E-cadherin dislocation in EFV-treated MCF10AT cells indicated the loss of cell-cell adhesion since these two proteins are typical adhesion molecules. In addition, western blot data confirmed the increased expression of the fibronectin and SLUG proteins in both EFV-treated MCF10AT and MCF10CA1α cells (Figure 3-13). Moreover, western blot data further illustrated decreasing CDC42, a crucial protein for cell skeleton formation, in EFV-treated MCF10AT and MCF10CA1α cells, suggesting that the cell skeletons might be dysregulated when levels of CDC42 are low. All these changes could be signs of EMT, metastasis, and invasiveness. They may be linked to cancer malignancy and poor therapeutic outcomes. However, at the same time, EFV-treated cells expressed more E-cadherin, a classic epithelial marker, in the same batch of samples. Similarly, less DDX5, a protein related to cancer malignancy, was detected in EFV-treated cells compared with untreated controls (Figure 3-13). It appears that EFV-treated MCF10AT and MCF10CA1α cells expressed some epithelial markers and some mesenchymal markers, at the same time, emphasising the complexity of the drug's effects on TNBCs.

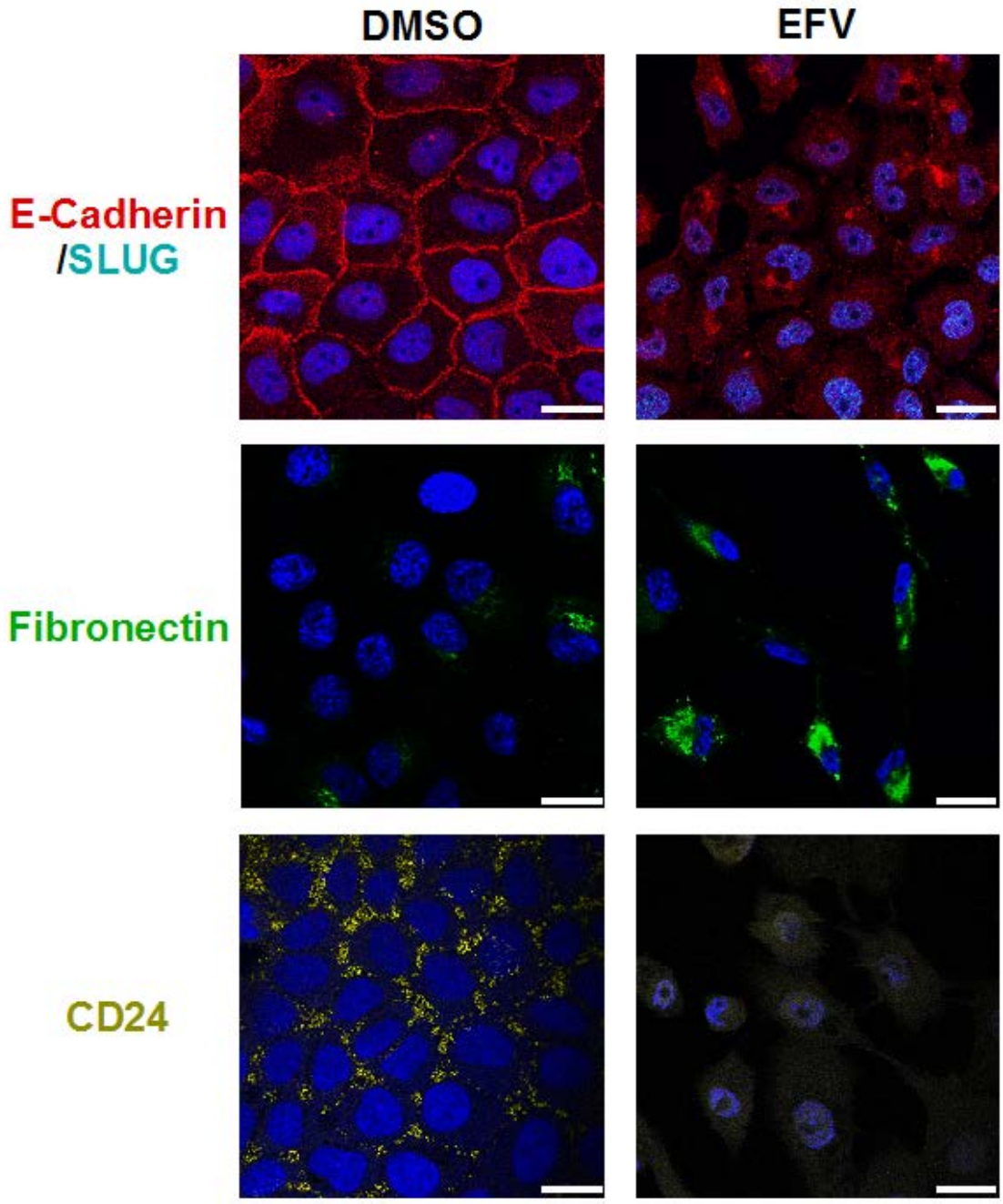


Figure 3–12 Immunofluorescence experiments for detecting the protein expression patterns of various morphological markers in untreated-control and EFV-treated MCF10AT cells
The untreated-control and EFV-treated MCF10AT cells were stained with different antibodies, representing various cell markers including E-cadherin (red), SLUG (light blue), fibronectin (green), CD24 (yellow), and nucleus (blue). (Scale bar: 25 µm)

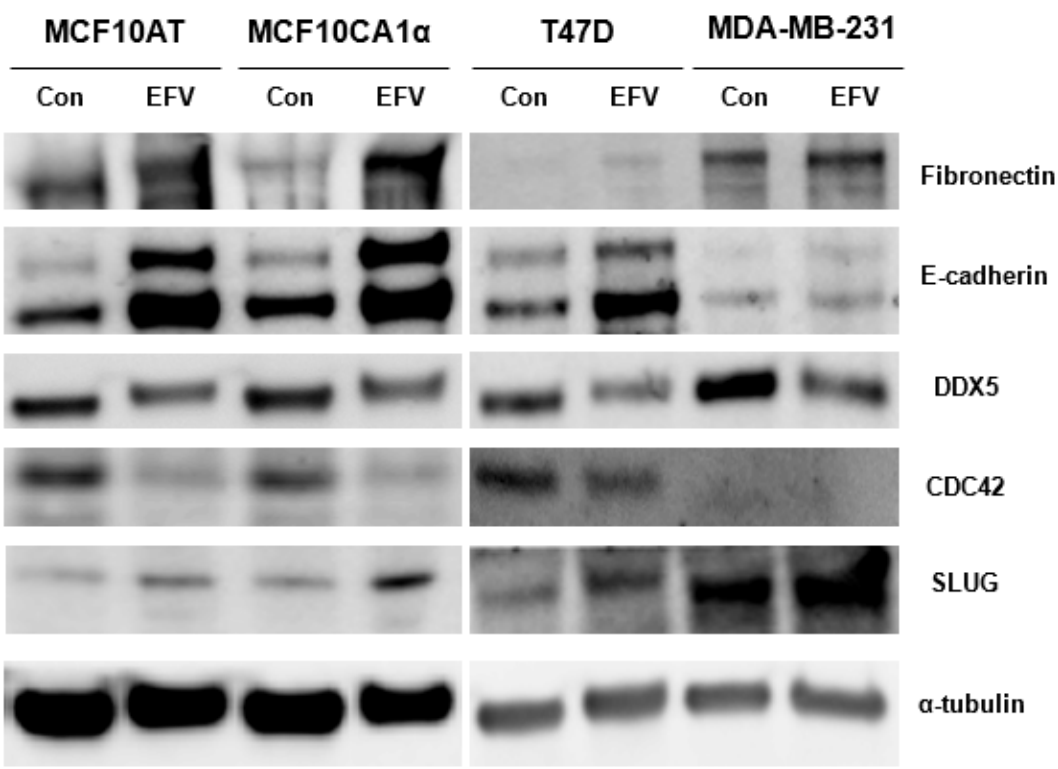


Figure 3–13 Expression patterns of cell morphology proteins in EFV-treated cells and untreated-control cells

The untreated-control and EFV-treated MCF10AT, MCF10CA1α, MDA-MB-231, and T47D cells were detected by western blotting with different antibodies representing various cell markers including fibronectin, E-cadherin, DDX5, CDC43, SLUG, and α-tubulin as internal control.

3.2.7 Antiretroviral drug treatment in other TNBC cell lines

TNBC heterogeneity may be responsible for the inconsistent drug responses observed with antiretroviral drugs. Breast cancer is heterogeneous, with morphological variation and molecular alteration able to be detected in different regions of a tumour (Fisher et al., 2013). Histologic and molecular heterogeneity have been widely observed in TNBCs and are considered one of the main obstacles to anti-TNBC drug development (Abramson and Mayer, 2014, Mills et al., 2018). In previous experiments, MCF10AT and MCF10CA1α showed a few distinct patterns with the MDA-MB-231 cell line suggesting diverse drug responses were displayed in various TNBC cell lines. To determine the consistency of antiretroviral drug responses on TNBC cell lines, two other commonly used TNBC cell lines, BT-549 and BT-20, were chosen to investigate antiretroviral drug responses in TNBCs.

Some experiments which were similar to the drug response experiments conducted on MCF10AT, MCF10CA1α, and MDA-MB-231 had been studied in BT-549 and BT-20 cell lines. These included cell viability, cell proliferation, cell cycle, and cell morphological experiments. Most of the experimental methods were consistent with previous experiments; except for the cell proliferation experiment which used immunofluorescence to detect the expression of the Ki-67 protein to indicate cell proliferation, in the cell nucleus. These experiments might be able to illustrate a greater view of antiretroviral drug responses on TNBCs and to further confirm the potential of antiretroviral drugs to treat TNBCs effectively.

3.2.7.1 *Anticancer capacities of EFV and SPV in BT-20 and BT-549 cell lines*

A cell viability assay was conducted to determine the correct drug concentrations for treating BT-20 and BT-549 TNBC cells (Figure 3-14), and the EC50 values of EFV and SPV in BT-20 and BT-549 cell lines were calculated by Prism software (Table 3-3), as described previously. The EC50 values of both EFV and SPV for BT-20 were lower than the EC50 values of BT-549, they were 18.56 µM for EFV and 19.25 µM for SPV. While, in BT-549 cells, the EC50 value for EFV was 30.55 µM and for SPV was 51.94 µM. This experiment suggested that BT-20 could be less tolerant of antiretroviral drugs than other TNBCs including MCF10AT, MCF10CA1α, MDA-MB-231, and BT-549. Among all the tested TNBC cell lines, BT-549 cell line needed the highest antiretroviral drug concentration for reaching the EC50 points.

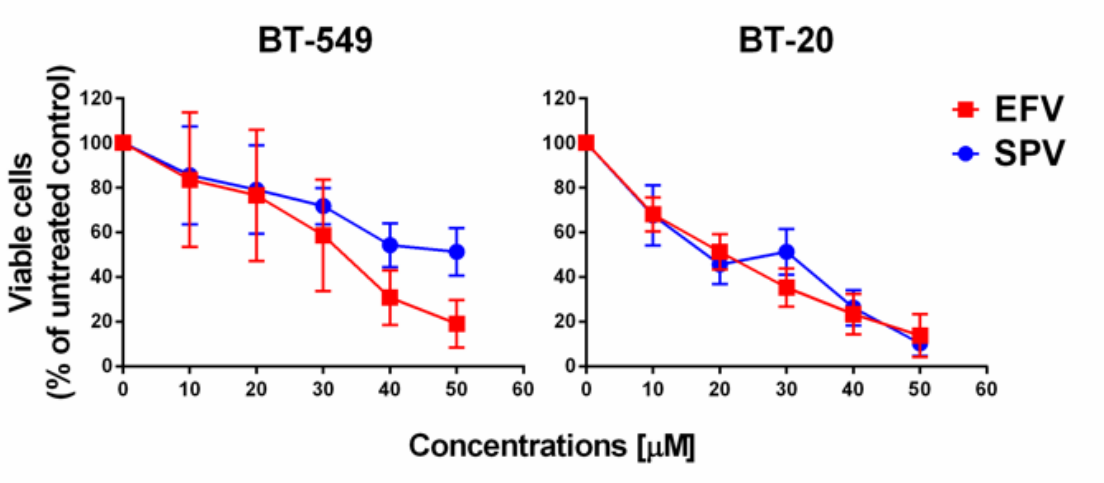


Figure 3–14 Cell viabilities when exposed to different antiretroviral drug concentrations
Cell viability curves of EFV treatment (red) and SPV treatment (blue) in BT-549 and BT-20 cells. Cell viability was determined at day 4 using a cell viability assay. Error bars: ±SD, n = 3.

Table 3-3 The EC50 values of EFV and SPV in BT-549 and BT-20 cell lines

	BT-549	BT-20
EFV	30.55 µM	18.56 µM
SPV	51.94 µM	19.25 µM

3.2.7.2 LINE-1 drug inhibition in BT-20 and BT-549

After determining the EC50 values for treating BT-549 and BT-20, the efficiency of LINE-1 inhibition by EFV in BT-20 and BT-549 was examined. Western blot data (Figure 3-15) were consistent with the previous finding that the expressions of LINE-1 ORF1 and LINE-1 RT can be reduced by EFV treatment in both BT-20 and BT-549 cell lines.

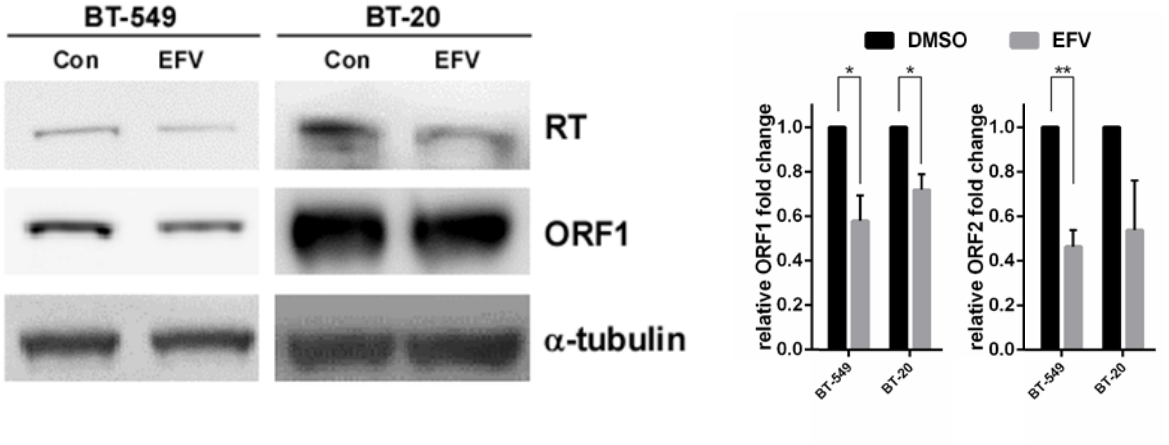


Figure 3–15 LINE-1 protein expression in BT-549 and BT-20 cell lines
The western blot data showed that LINE-1 is expressed in BT-549 and BT-20 cell lines. The relative ORF1 and ORF2 RT expressions were reduced after EFV treatment. α-tubulin was an internal control. Both ORF1 and ORF2 RT expressions decreased after EFV treatment in BT-549 and BT-20. **p*-value <0.05; ***p*-value <0.01. Error bars: ±SD, n = 3.

3.2.7.3 The effects of EFV on cell proliferation and cell cycle in BT-20 and BT-549

After EFV was confirmed to target BT-20 and BT-549 cells and to inhibit LINE-1 expression in BT-20 and BT-549, its precise effects were investigated in BT-20 and BT-549 cells.

Cell cycle data for BT-20 and BT-549 cells (Figure 3-16 and in appendix Figure S3-6) were consistent with the other tested cancer cell lines. EFV treatment reduced the percentage of cells in S phase (BT-549: 19.46±3.46% to 15.70±0.17%; BT-20: 12.61±2.35% to 9.86±1.51%) and increased the percentage of cells in G2 phase compared with their untreated control (BT-549: 21.70±2.17% to 25.80±4.32%; BT-20: 25.52±3.33% to 36.66±5.05%).

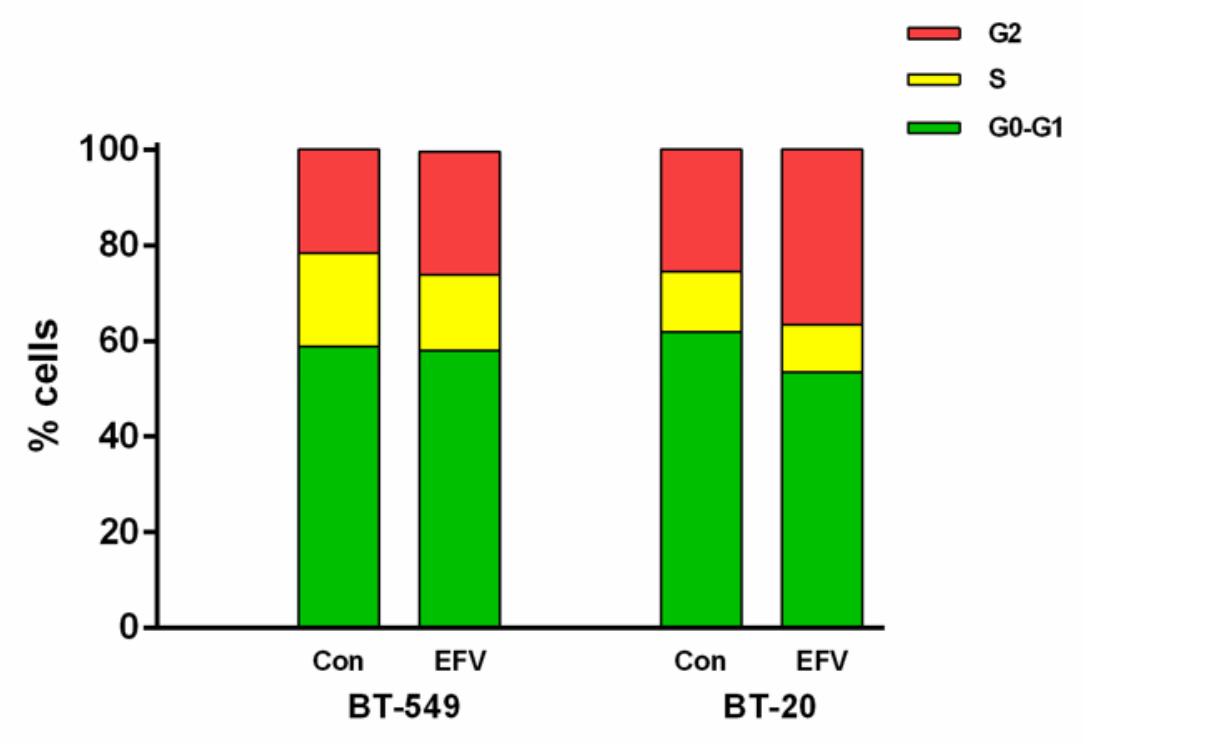


Figure 3–16 Cell cycles of untreated-control and EFV-treated cells in BT-20 and BT-549 cell lines
The percentage of G2 phase cells increased after EFV treatment, and the percentage of S phase cells decreased after EFV treatment (n = 3).

Drug responses of BT-549 and BT-20 were observed via immunofluorescence experiments. Ki-67 (a cell proliferation marker) staining in antiretroviral drug-treated and untreated BT-20 and BT-549 cell lines (Figure 3-17) indicated that antiretroviral drug treatment is able to reduce cell proliferation in these two cell lines. Previous observations showed fewer Ki-67-stained-red dots in antiretroviral drug-treated BT-20 and BT-549 cells, compared with their untreated-controls. Also, phalloidin cell morphology staining (Figure 3-17) showed very clear morphological differences between antiretroviral drug-treated and untreated cells indicating that antiretroviral drug-treated BT-549 and BT-20 cells were differentiated because of distinct cell projection structures (more images in Appendix Figure S3-7). These responses were similar to MDA-MB-231 cells and not MCF10AT and MCF10CA1α cells, suggesting that the drug effects of BT-549 and BT-20 might be close to the effects of MDA-MB-231; while MCF10AT and MCF10CA1α could have slightly different drug responses to other TNBCs. This suggested that drug responses could be varied because of the diverse genetic background in TNBCs.

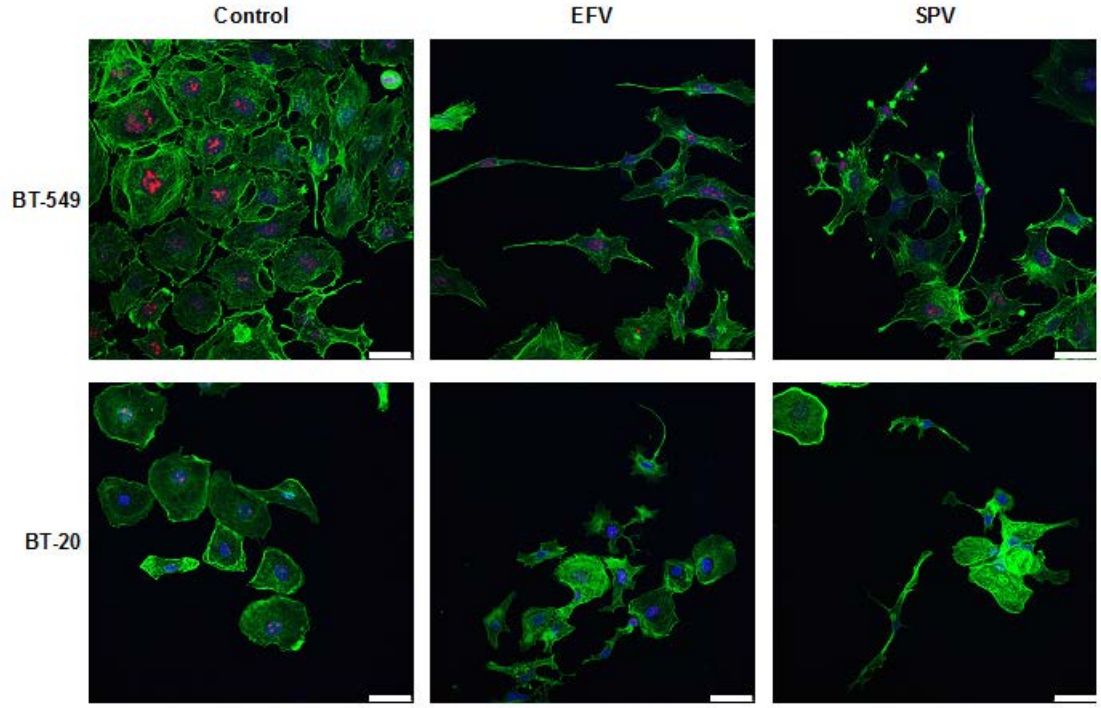


Figure 3–17 Cell morphologies in EFV- and SPV-treated and untreated cells
This figure compares the morphologies of antiretroviral drug-treated and untreated BT-549 and BT-20 cell lines. Phalloidin (green, for detecting F-actin), Ki-67 (red, for detecting cell proliferation), and DAPI (blue, for detecting nucleus) were used for immunofluorescence staining (Scale bar: 50 μm).

3.3 Discussion

According to two different papers published in 2005, Efavirenz treatment can effectively reduce cell numbers to less than half of untreated controls and enhance G0-G1 phase cells but decline S phase cells in human thyroid carcinoma (ARO and FRO), prostate carcinoma (PC3), and melanoma (A-375) cell lines (Landriscina et al., 2005; Sciamanna et al., 2005). This anti-proliferation effect has also been reported in the non-TNBC T47D cell line and pancreatic cancer BxPC-3 and Panc-1 cell lines (Ranala et al., 2014; Hecht et al., 2015). Hecht's team specifically pointed out the cytotoxic effect (by inducing apoptosis and necrosis) of Efavirenz implying its potential use as an anti-cancer drug. Our TNBC data is highly consistent with all these findings apart from a minor difference where cells in the G0-G1 phase show an increase in MCF10AT and MCF10CA1α cell lines but a decrease in MDA-MB-231, BT-549, and BT-20 cell lines. This suggests Efavirenz could act as a potential anti-TNBC drug with systemic effects in diverse cancers.

In this chapter, we have shown the effectiveness of EFV and SPV at eliminating TNBC cells but not non-cancerous MCF10A cells. We have also demonstrated the effectiveness of EFV and SPV at reducing reverse transcription activities of LINE-1. EFV and SPV were then used for the remaining experiments presented in this thesis. The results implied a potential link between LINE-1 inhibition and breast cancer elimination when using antiretroviral drug treatment, which was consistent with previous research in other cancers and non-TNBCs. Further studies were conducted to confirm that antiretroviral drugs can retard cell division, reduce DNA synthesis within the cell cycle, and induce programmed and non-programmed cell death in TNBCs. Moreover, clear morphological changes are observed in the drug-treated cancer cells. These results strongly suggested that antiretroviral drugs were functioning by inducing TNBC cell alterations. One of the critical challenges we encountered when conducting these experiments was LINE-1 overexpression and inhibition. To overexpress LINE-1 in non-cancer cells is difficult because there are various self-defence mechanisms, including methylation and degradation present in normal cells to prevent LINE-1 expression and protect cell status. Thus, the plasmid with the modified LINE-1 sequence, instead of the original LINE-1 sequence, was utilized for the experiments. Moreover, many commonly used non-cancerous breast cells with very limited LINE-1 expression, such as MCF10A and primary mammary epithelial cells (HMEC), might be insufficiently competent making it difficult to insert plasmid

DNA into these cells. Therefore, we used the same batch of transfected cells for the LINE-1 overexpression experiment rather than conducting three independent transfections. On the other hand, LINE-1 inhibition by shRNA was more challenging than LINE-1 overexpression in TNBCs. Thus, the LINE-1 silencing plasmid was previously used in the T47D (luminal-type breast cancer) cell line to investigate the responses of LINE-1 direct inhibition (Chen et al., 2012). The results suggested that direct inhibition of LINE-1 by shRNA gave similar responses to inhibition of LINE-1 by EFV. This result confirmed that EFV treatment might inhibit LINE-1 in the same way that RNA did. However, silencing of LINE-1 shRNA had very limited effects in MCF10AT and MCF10CA1α cell lines. One reason for this could be that the cell proliferation rates of MCF10AT and MCF10CA1α cells were much greater than T47D cells. Consequently, the shRNA cannot appropriately inhibit LINE-1 expression. A slow-growing TNBC may be able to improve the observation of shRNA inhibition effects. Yet, most of the well-studied TNBC cell lines proliferate faster than other breast cancer cell lines.

Inconsistent drug-induced morphological changes, possibly caused by TNBC heterogeneity, were equally challenging. Much evidence including immunofluorescence and western blot experiments supported previous observations that antiretroviral drugs were able to change the status of TNBCs. However, we have presented controversial data in disagreement with the original hypothesis i.e. antiretroviral drugs can return cancer cells back into normal-like cells, in MCF10AT and MCF10CA1α cell lines. Although there is no doubt that antiretroviral drugs induce cell differentiation in TNBC cell lines; the phenotypes which are associated with cancer invasion, metastasis, and recurrence were also observed at the same time. The detection of molecular markers after the treatment further suggested the drug treatments may lead to alternative prognostic outcomes in TNBCs.

There are two major limitations worth discussing: Firstly, many TNBC cells express limited E-cadherin protein and some breast cancer cells lack fibronectin expression endogenously. Therefore, it could be difficult to compare changes in the levels of these proteins. Secondly, inconsistencies in the drug responses among MCF10AT and MCF10 CA1α and MDA-MB-231 cell lines led us to conduct similar experiments in BT-549 and BT-20 cell lines. We did this in order to confirm which protein and molecular changes could be universal responses to antiretroviral drugs by TNBCs. The results suggested that BT-549 and BT-20 cell lines respond

similarly to the MDA-MB-231 cell line and less so to the MCF10AT and MCF10CA1α cell lines. Limited time and resources restricted the number of cell lines that we were able to test, hence our conclusions are based on the pattern of drug responses in these five TNBC cell lines.

The diversity and heterogeneity of TNBCs also need to be considered. Recently, it has been proposed that TNBCs should be separated into the following subtypes, *basal-like 1* subtype, *basal-like 2* subtype, *immunomodulatory* subtype, *mesenchymal-like* subtype, *mesenchymal stem-like* subtype, *luminal AR* subtype, and others (Lehmann et al., 2011). This new classification method highlighted that breast cancer cells with $ER^-/PR^-/HER^-$ features are all classified as TNBCs, but might have different drug responses as a result of the variation in their genetic backgrounds. For example, some of them are highly enriched in cell cycle-related pathways like *basal-like 1* and *2* subtypes; and others strongly express pathways relevant to cell motility and cell differentiation, such as *mesenchymal-like* and *mesenchymal stem-like* subtypes (Lehmann et al., 2011). Each of these different subtypes of TNBCs could have diverse drug responses and might need to be treated differently.

To further understand how these TNBC cells were affected by antiretroviral drugs, we investigated their gene regulation. We used next-generation RNA sequencing (RNA-Seq) to compare the differences in RNA level between antiretroviral drug-treated cells and untreated-control cells. It was hoped that RNA-Seq analysis might help to explain why some mesenchymal markers were unexpectedly detected in antiretroviral drugs-treated cells. We were also interested to see whether the gene regulation data would agree with the western blot data, shown previously. The details are outlined in the next chapter which mainly focuses on how antiretroviral drugs regulate the anticancer process, and how we can improve the strategies for treatment by identifying the key genes/pathways involved in the anti-cancer process.

Chapter 4 Anticancer mechanisms of antiretroviral drugs in triple-negative breast cancers

4.1 Introduction

The main theme explored in this study was whether antiretroviral drugs can act as anticancer agents in triple-negative breast cancers (TNBCs). In the previous chapter, antiretroviral drugs Efavirenz (EFV) and SPV122 (SPV) were found to decrease cancer cell proliferation and to promote cancer cell death, thus effectively reducing breast cancer cells. In parallel, LINE-1 expression was suppressed by treatment with antiretroviral drugs in all tested breast cancer cell lines. Since LINE-1 expression could be one of the causes of cancer development (Sciamanna et al., 2018), antiretroviral drugs may effectively antagonize TNBC growth through LINE-1 inhibition. However, antiretroviral drug-treated MCF10AT and MCF10CA1α TNBC cell lines were found to display some unexpected mesenchymal markers indicating unfavourable prognostic outcomes such as cancer metastasis and cancer recurrence. These findings are at odds with our original hypothesis and further emphasise the complexities of TNBCs.

The second goal in this study was to determine the mechanisms which govern the anticancer properties of antiretroviral drugs in TNBCs. Despite the fact that some antiretroviral drugs with potential for treating certain cancers are now in clinical trials (Sciamanna et al., 2018), the molecular and pathway regulations remain vague. RNA sequencing (RNA-Seq) is a powerful technique for studying epigenetic regulation. It helps to identify the phenotypic variations among different treatment conditions (Evsikov and Marin de Evsikova, 2016). In this chapter, some epigenetic experiments were conducted, mainly based on RNA-seq analysis, which improves the understanding of the gene and pathway regulations for antiretroviral drug in cancer cells and to explore possible anticancer mechanisms of this class of drug.

Here, the basic process of RNA-Seq analysis, including experimental design, data analysis methods are described. This is followed by a discussion of the gene and pathway regulations involved in antiretroviral drug treatments. This also includes analysis of data which at first glance contradicts the original predictions that were first made. Finally, the potential mechanisms behind antiretroviral drug-induced cancer responses are discussed, as are the potential roles of cancer stem cells (CSCs) in the process.

4.2 Results

4.2.1 The process of RNA sequencing analysis

RNA sequencing (RNA-Seq) is capable of generating massive data for studying the transcriptome; therefore, it is important to interpret all information correctly and accurately. Below, the analytical methods which were used to identify important genes and pathways from the RNA-Seq data were briefly described.

4.2.1.1 Analysis of variance (ANOVA)

Analysis of variance (ANOVA) is one of the most widely used statistical methods for analysing RNA-Seq data. As reported in the RNA-Seq results section, some genes showed dramatic changes after antiretroviral drug treatment. ANOVA analysis compares group means, enhances analysis and highlights the significance of these changes. *q*-values, instead of the traditional *p*-values, are calculated to indicate whether each gene shows a statistically significant variation after antiretroviral drug treatment. The use of the *p*-value as a criterion in a multiple testing analysis, like RNA-Seq, may result in an increase in the false-positive rate with the increased numbers of samples (lose the power) in multiple tests (Storey, 2003). Unlike *p*-values, *q*-values (optimised *p*-values) take into consideration the false discovery rate (FDR) and have a higher chance of detecting significant differences (Li et al., 2007). After the FDR correction, the appropriate cut-off for genes with significant differences in this RNA-Seq analysis was set as the *q*-value less than 0.25 (Yu et al., 2012) (Table 4-1). These genes are likely to have a role in antiretroviral drug-induced cancer responses.

Table 4-1 The numbers of genes in different *q*-value cut-offs in ANOVA analysed RNA-Seq result

	<0.001	<0.01	<0.025	<0.05	<0.1	<0.2	<0.25	<0.275	<0.3	<0.4	<1
EFV	0	0	0	0	0	371	848	981	1284	2918	11282
SPV	0	0	0	0	0	778	1465	1860	2294	3886	11282

Genes associated with drug responses may vary depending upon different experimental conditions and the mechanism of the drugs. The top ten candidate genes, displaying the greatest changes according to the RNA-Seq analysis upon EFV and SPV treatment, are exhibited in Table 4-2. Some candidate genes are in both drug lists; whereas, others are listed in either the EFV-treatment list or the SPV-treatment list. Candidate genes such as stearoyl-CoA desaturase (*SCD*), alkaline phosphatase, placental (*ALPP*), and ETS-related transcription factor 3 (*ELF3*) showed extensive changes after treatment with antiretroviral drugs. Another candidate genes including pyruvate dehydrogenase kinase, isozyme 4 (*PDK4*), Guanylate-binding protein 7 (*GBP7*), sterol regulatory element-binding transcription factor 1 (*SREBF1*), complement component 3 (*C3*), acyl-CoA synthetase long-chain family member 5 (*ACSL5*), and high mobility group AT-hook 2 (*HMGA2*) were found among the EFV-treatment top ten. Other candidate genes including interferon-induced protein with tetratricopeptide repeats 3 (*IFIT3*), angiopoietin-like 4 (*ANGPTL4*), retinoic acid receptor responder 1 (*RARRES1*), IGF-like family member 1 (IGFL1), kallikrein-related peptidase 10 (*KLK10*), *CD14*, and interferon-induced protein with tetratricopeptide repeats 2 (*IFIT2*) were among the SPV-treatment top ten. Interestingly, the expression of overlapped candidate genes including *SCD*, *ALPP*, and *ELF3* decreased after antiretroviral drug treatment which may potentially be linked to better prognostic outcomes. Further analyses were conducted based on this preliminary ANOVA analysis and were discussed later on.

Table 4-2 The top 10 candidate genes involved in antiretroviral drug-induced cancer responses

Gene name	*Drug response*	*Role in cancers*	*References*
		EFV or SPV-treated TNBCs	
SCD stearoyl-CoA desaturase	↓	Cancer proliferation, tumour expansion (prostate, breast cancer)	(Kim et al., 2011, Belkaid et al., 2015, Peck et al., 2016)
ALPP alkaline phosphatase, placental	↓	Tumour marker (seminoma, ovarian, and metastatic prostate cancer)	(Fishman, 1987, Albrecht et al., 2004, Rao et al., 2016)
ELF3 ETS-related transcription factor 3	↓	Metastasis (colorectal and lung cancer), poor prognosis (colorectal)	(Nakarai et al., 2012, Wang et al., 2014b, Wang et al., 2018)
		EFV-treated TNBCs	
PDK4 pyruvate dehydrogenase kinase, isozyme 4	↓	Tumorigenesis, drug resistance (colon, breast, lung and pancreatic cancer)	(Liu et al., 2014b, Leclerc et al., 2017, Woolbright et al., 2018)
GBP7 Guanylate-binding protein 7	↓	Unknown	
SREBF1 sterol regulatory element-binding transcription factor 1	↓	Cancer cell growth (prostate and colon cancer), tumorigenesis (pancreatic cancer)	(Griffiths et al., 2013, Sun et al., 2015, Audet-Walsh et al., 2018, Wen et al., 2018)
C3 complement component 3	↓	Epithelial cancer marker and EMT marker (ovarian & endometrial cancer)	(Cho et al., 2016)
ACSL5 acyl-CoA synthetase long-chain family member 5	↓	Cancer cell survival (glioma cell)	(Mashima et al., 2008)
HMGA2 high mobility group AT-hook 2	↑	Cancer proliferation and metastasis (epithelial cancer)	(Morishita et al., 2013, Wu et al., 2016b, Gao et al., 2017)
		SPV-treated TNBCs	
IFIT3 interferon-induced protein with tetratricopeptide repeats 3	↓	Poor prognostic outcome (pancreatic ductal adenocarcinoma)	(Zhao et al., 2017)
ANGPTL4 angiopoietin-like 4	↑	Tumorigenesis	(Tan et al., 2012)
RARRES1 retinoic acid receptor responder 1	↓	Tumour suppressor, autophagy inducer and angiogenesis inhibitor (TNBC)	(Sahab et al., 2011)
IGFL1 IGF-like family member 1	↓	CSC marker	(Birnie et al., 2008)
KLK10 kallikrein-related peptidase 10	↓	Cell proliferation repressor, apoptosis inducer (oesophageal and prostate cancer)	(Hu et al., 2015b, Li et al., 2015)
CD14 CD14 molecule	↓	Tumour growth, EMT (gastric cancer)	(Cheah et al., 2015)
IFIT2 interferon-induced protein with tetratricopeptide repeats 2	↓	Prevents metastasis (oral squamous cell carcinoma)	(Lai et al., 2012)

(The gene changes related to positive outcomes are labelled with a yellow background; the gene changes related to negative outcomes are labelled with the grey-blue background. EMT: epithelial-to-mesenchymal transition. CSC: cancer stem cell.)

4.2.1.2 STRING-DB analysis

Antiretroviral drugs affect certain genes through various pathways in TNBCs. To further investigate whether the responses of TNBC cells to such drugs were associated with any existing signalling pathway, the STRING-DB pathway search tool (version: 10.5, https://string-db.org/) (Szklarczyk et al., 2017) was used. Genes in the same signalling pathways may be up regulated or down regulated according to their roles. However, because of its algorithm, STRING-DB imports too many genes at once may sometimes conceal some potential pathways. Therefore careful information input is critical for the STRING-DB analysis. Six gene sets: EFV up regulated, EFV down regulated, EFV up and down, SPV up regulated, SPV down regulated, and SPV up and down, was established based on different selection criteria. Firstly, all the candidate genes had to show significant drug-induced differences (*q*-value less than 0.25). Additionally, more weight was placed on the genes displaying greater response differences after treatment. Only the genes with higher variations were included (up regulated: fold-change greater than 1.2, or down regulated: fold-change less than -1.2). These gene sets were uploaded to the STRING-DB online database to analyse the potential protein-protein association networks. From this, it was possible to determine signalling pathways involved in antiretroviral drug treatments, based on the Kyoto Encyclopaedia of Genes and Genomes (KEGG) database (https://www.genome.jp/kegg/) and the Gene Ontology Consortium database (http://www.geneontology.org/). More details are discussed later in this chapter.

4.2.1.3 Network analysis

The RNA-Seq data included a massive amount of information, and one runs the risk of losing important details by only using ANOVA analysis. Network analysis is an alternative method of analysis that overcomes this problem by grouping genes with similar drug responses. The concept of network analysis is that the genes exhibiting similar expression patterns, under different conditions (cell lines and treatments), may be related and belong to corresponding signalling pathways. Thus, network analysis was utilised to investigate whether any significant pathway could be emphasised to reinforce the analysis of RNA-Seq.

In network analysis, many separate modules representing different gene expression patterns were observed. Although there were forty-four modules identified using this approach, only eleven passed the selection criteria (*q*-value less than 0.05) (Figure 4-1) and were regarded as

valid. The genes in these modules were analysed by the STRING-DB search tool to identify the significant pathways involved.

In the original experimental design of the RNA-Seq for this project, gene regulations in the DMSO-control cells with antiretroviral drug-treated cells in MCF10A, MCF10AT, MCF10CA1α, MDA-MB-231, and MCF7 cell lines were compared. In the initial analysis, only the gene regulations in TNBC cell lines were analysed because the MCF7 cell line was considered to be a non-TNBC luminal-type breast cancer control. Due to the expression of its oestrogen receptors and progesterone receptors, the MCF7 cell line may launch alternative signalling pathways and gene regulations than the TNBCs. However, candidate genes regulated by drug treatments in the TNBC-only gene set and the TNBC+MCF7 gene set were partially identical. Moreover, the candidate gene with the greatest fold-change, *SCD*, was identical in the TNBC-only and the TNBC+MCF7 gene set. Also, the significant signalling pathways highlighted by the STRING-DB search tool overlapped in both EFV and SPV ANOVA lists. These findings suggested that the anticancer effects of antiretroviral drugs in TNBCs may share some mechanisms with MCF7. In addition, the more samples imported to the network analysis, the easier it is to find genes with similar patterns owing to greater tolerance in a larger dataset. Therefore, the network analysis was done for both TNBC-only and TNBC+MCF7 gene sets. The candidate pathways involved in antiretroviral drug treatments are discussed in the following sections.

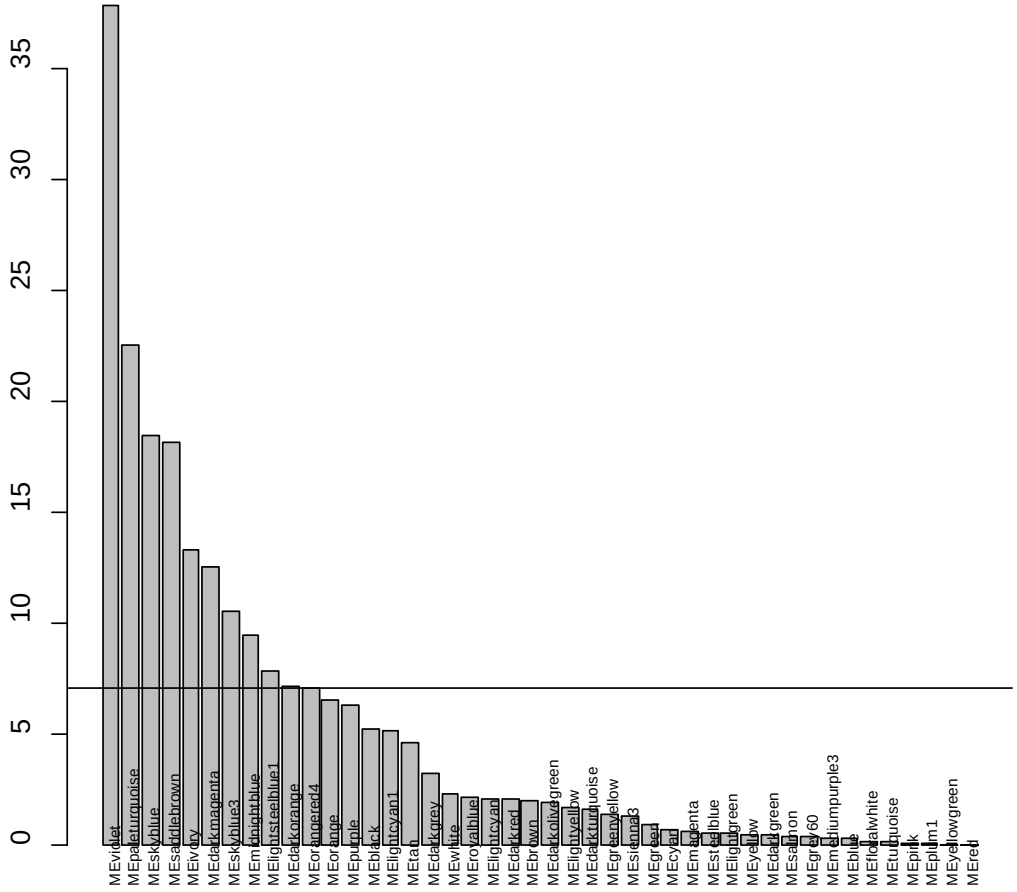


Figure 4-1 Bar Plot of ANOVA *q*-values for module eigengenes in the network analysis
The modules identified by network analysis are listed on the x-axis. The horizontal line is the 0.05 cut-off for the Storey *q*-values. The nine modules which pass the cut-off are considered as significant modules.

4.2.2 Down regulated pathways after antiretroviral drug treatment

As mentioned above, after ANOVA analysis, STRING-DB analysis was used to further discover relevant genes and pathways which may be affected by antiretroviral drugs in TNBCs. Combination of down regulated and up regulated gene sets, which included more than 400 genes, may reduce the number of significant results due to the algorithms of STRING-DB. Thus, the first step was to focus on possible genes and pathways involved in the down regulation of genes by antiretroviral drugs. The results of the EFV down regulated gene set and the SPV down regulated gene set were discussed separately, since the genes listed in these two gene sets could be different due to the pharmacological mechanisms of the drugs.

Three significant KEGG pathways were highlighted in the EFV down regulated gene set. This gene set contained 237 genes (in appendix Table S4-1) that fitted the selection criteria (*q*-value less than 0.25, fold-change less than -1.2). These genes were imported and analysed by the STRING-DB to discover potential genes and pathways involved in the EFV-induced cancer responses. 231/237 genes had been identified their protein products, and 232 potential links had been calculated among these proteins. The number of predicted links were greater than the number of expected links (141 links) and the protein-protein interaction (PPI) enrichment *p*-value was less than 10^{-16} suggesting that these proteins might have some biological connections. The significant KEGG pathways (Figure 4-2 and Table 4-3) arising from this gene set were 'fatty acid metabolism' (seven genes, *q*-value = 0.0003), 'biosynthesis of unsaturated fatty acids' (four genes, *q*-value = 0.0078) and 'FoxO signalling pathway' (eight genes, *q*-value = 0.0078). Only four candidate genes were involved in 'biosynthesis of unsaturated fatty acids', and three of them were shown in 'fatty acid metabolism' pathways. Thus 'biosynthesis of unsaturated fatty acids' could be considered a relevant branch of 'fatty acid metabolism'. Consequently, only 'fatty acid metabolism' and 'FoxO signalling pathway' were focused on in this thesis and are discussed later in the chapter.

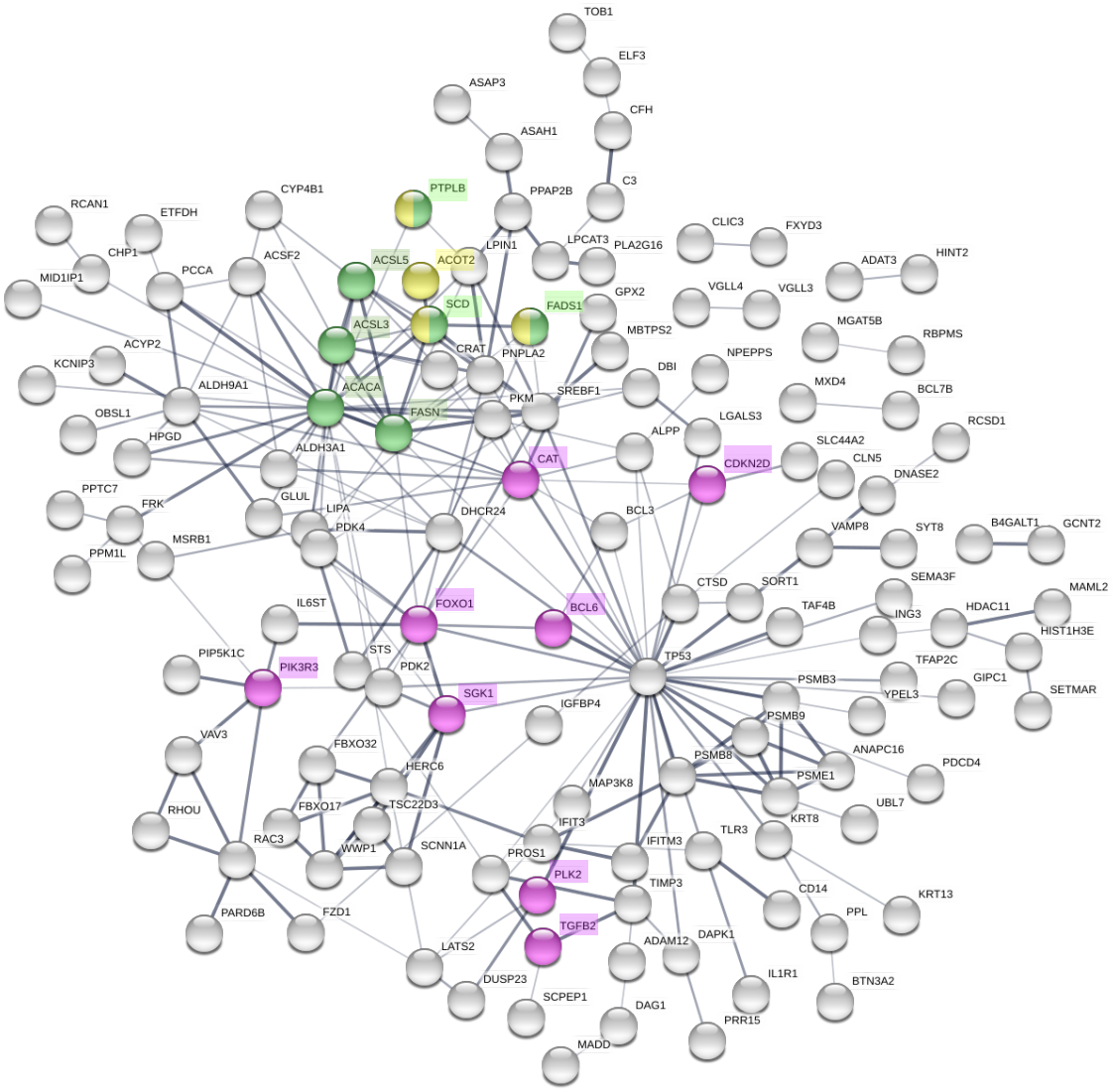


Figure 4-2 The STRING-DB gene network for EFV-treatment induced down regulated genes in TNBCs

This figure represents possible protein-protein interaction which was calculated by the STRING-DB software. A node represents a protein; a line represents a potential link. The thickness of the line indicates the strength of data support. The finest line represents medium confidence (0.400); the middle thickness line represents high confidence (0.700); the thickest line represents the highest confidence (0.900). The green nodes are proteins related to 'fatty acid metabolism'; the yellow nodes are proteins related to 'biosynthesis of unsaturated fatty acids'; the pink nodes are proteins related to 'FoxO signalling pathway'.

Table 4-3 Highlighted KEGG pathways in EFV down regulated gene set by STRING-DB

Pathway ID	Pathway description	Count in network	*q*-value
01212	**Fatty acid metabolism**	7	0.0003
01040	**Biosynthesis of unsaturated fatty acids**	4	0.0078
04068	**FoxO signalling pathway**	8	0.0078

Four significant KEGG pathways were highlighted in the SPV down regulated gene set. This gene set contained 379 genes (in appendix Table S4-2) which corresponded to the selection criteria (*q*-value less than 0.25 and fold-change less than -1.2). Not surprisingly, as more genes were imported into the STRING-DB search tool, more pathways arose. Four significant pathways were highlighted in the SPV down regulated gene set; whereas, only three significant pathways were listed in the EFV down regulated gene set. 374/379 genes had been identified their protein products, and 860 potential links had been calculated among these proteins. The number of predicted links were greater than the number of expected links (552 links) and the PPI enrichment *p*-value was less than 10^{-16} suggesting that these proteins might have some biological connections. The most significant KEGG pathway from the SPV down regulated gene set (Figure 4-3 and Table 4-4) was 'fatty acid metabolism' (eleven genes, *q*-value = 2.33×10^{-7}) which was similar to the result of the EFV down regulated gene set. These results highlight the importance of 'fatty acid metabolism' in antiretroviral drug-induced cancer cell responses. Many genes in the other highlighted pathways were also shown in 'fatty acid metabolism'. For instance, 'fatty acid degradation' (seven genes, *q*-value = 0.00169) had five genes, 'biosynthesis of unsaturated fatty acids' (five genes, *q*-value = 0.00304) had four genes, and 'PPAR signalling pathway' (seven genes, *q*-value = 0.0172) had five genes which overlapped with 'fatty acid metabolism' genes, respectively. This suggested these three pathways could be highly relevant to 'fatty acid metabolism'; therefore, 'fatty acid metabolism' was prioritised as the main focus of the SPV down regulated gene set.

As predicted, the most significant pathway in the EFV down regulated gene set and the SPV down regulated gene set was the same. This implied that EFV and SPV may share the same target and thus induce a similar pathway down regulation. At the same time, these results emphasised the importance of 'fatty acid metabolism' in antiretroviral drug treatment in TNBCs.

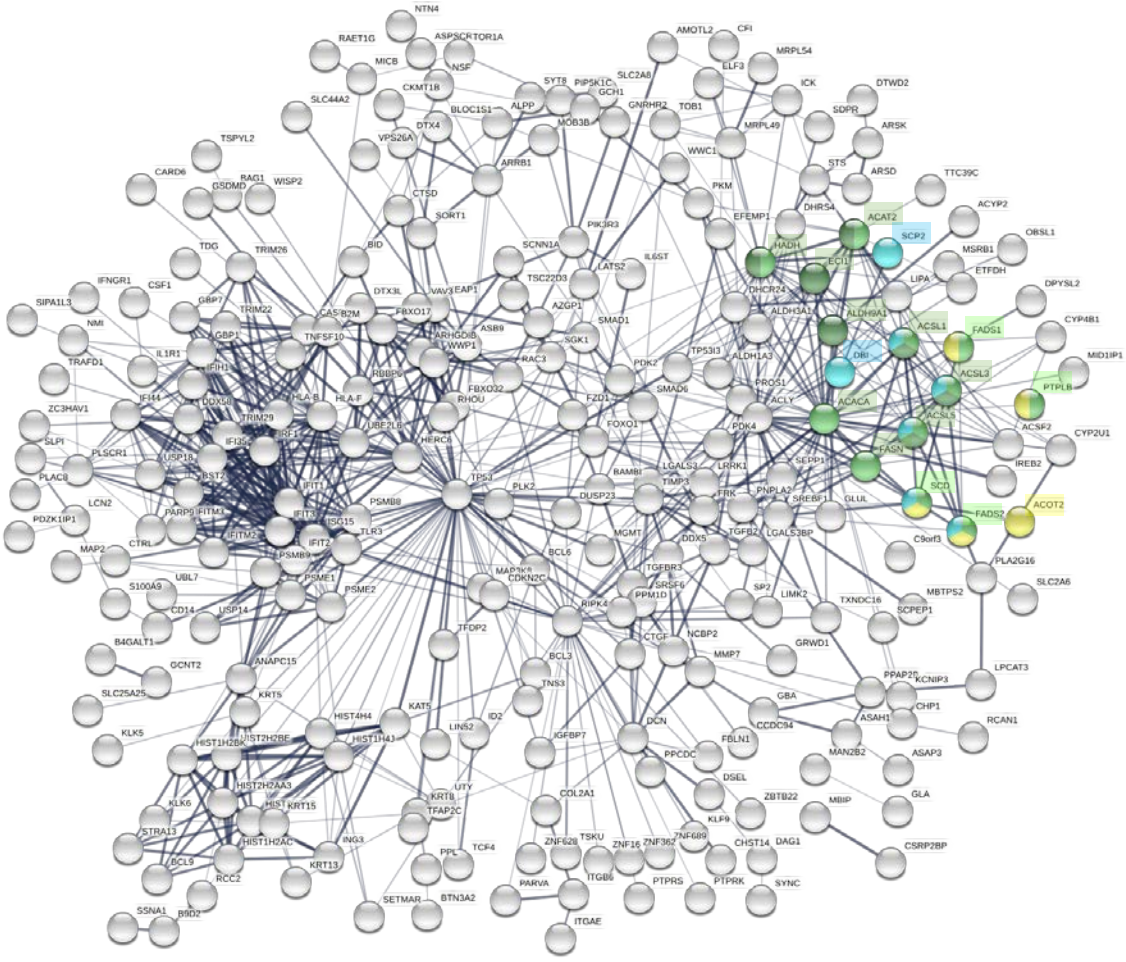


Figure 4-3 The STRING-DB gene network for SPV-treatment induced down regulated genes in TNBCs

This figure represents possible protein-protein interaction which was calculated by the STRING-DB software. A node represents a protein; a line represents a potential link. The thickness of the line indicates the strength of data support. The finest line represents medium confidence (0.400); the middle thickness line represents high confidence (0.700); the thickest line represents the highest confidence (0.900). The green nodes are proteins related to 'fatty acid metabolism'; the dark green nodes are proteins related to 'fatty acid degradation'; the yellow nodes are proteins related to 'biosynthesis of unsaturated fatty acids'; the light blue nodes are proteins related to 'PPAR signalling pathway'.

Table 4-4 Highlighted KEGG pathway in SPV down regulated gene set by STRING-DB

Pathway ID	Pathway description	Count network	*q*-value
01212	**Fatty acid metabolism**	11	2.33×10^{-7}
00071	**Fatty acid degradation**	7	0.00169
01040	**Biosynthesis of unsaturated fatty acids**	5	0.00304
03320	**PPAR signalling pathway**	7	0.0172

4.2.2.1 Fatty acid metabolism

The anticancer effects of antiretroviral drugs appeared to be strongly associated with down regulation of fatty acid metabolism in cancers. The down-regulation of fatty acid metabolism-related genes had been emphasised not only in the KEGG database (Table 4-3 and Table 4-4) but also in the GO database according to the calculation of the STRING-DB. Also, *SCD* which had the greatest fold-change after antiretroviral drug treatment in TNBCs, according to the ANOVA analysis, was a crucial gene in fatty acid metabolism. Moreover, the heat maps (Figure 4-4), which represent RNA expression levels across different treatments, illustrated that most of the key fatty acid metabolism-related genes, including *SCD*, acyl-CoA synthetase long-chain family member 1 (*ACSL1*), acyl-CoA synthetase long-chain family member 3 (*ACSL3*), acyl-CoA synthetase long-chain family member 5 (*ACSL5*), fatty acid synthase (*FASN*), fatty acid desaturase 1 (*FADS1*), fatty acid desaturase 2 (*FADS2*), 3-hydroxyacyl-CoA dehydratase 2 (*HACD2* or *PTPLB*), acetyl-CoA carboxylase alpha (*ACACA*), sterol regulatory element-binding transcription factor 1 (*SREBF1*), ATP-citrate synthase (*ACLY*), acetyl-CoA acetyltransferase 2 (*ACAT2*), and Hydroxyacyl-CoA dehydrogenase (*HADH*) were decreased (red) after antiretroviral drug treatments, compared with their controls, in all breast cancer cell lines. These results were in agreement with many previous studies suggesting that fatty acid metabolism plays a critical role in cancer (Yi et al., 2018, Wang et al., 2017c, Kuo and Ann, 2018).

The fatty acid metabolism-related genes (according to the KEGG database) highlighted in both of the EFV up regulated gene set and the SPV up regulated gene set are listed in Table 4-5 with their potential roles in cancer. These genes included *SCD*, *ACSL5*, *FASN*, *ACSL3*, *FADS1*, *PTPLB* and *ACACA*. The order of these genes was descending of the fold-changes. Most of them (*SCD*, *ACSL5*, *FASN*, *ACSL3*, and *ACACA*) have been linked to tumorigenesis and tumour growth in many types of cancer (Kim et al., 2011, Tirinato et al., 2017).

Table 4-5 Antiretroviral drugs induced down regulated fatty acid metabolism-related genes

Gene names	***Roles in cancers***	***References***
	EFV or SPV-treated TNBCs	
SCD stearoyl-CoA desaturase	Cancer proliferation, tumour expansion (prostate, breast cancer)	(Kim et al., 2011, Belkaid et al., 2015, Peck et al., 2016)
ACSL5 acyl-CoA synthetase long-chain family member 5	Cancer cell survival (glioma cell)	(Mashima et al., 2008)
FASN fatty acid synthase	Cell proliferation, cell survival, cell adhesion, migration, and invasion (most human carcinomas)	(Wang et al., 2017c, Cui et al., 2017, Kuo and Ann, 2018)
ACSL3 acyl-CoA synthetase long-chain family member 3	Tumorigenesis, cancer initiation (lung cancer)	(Perera et al., 2009, Padanad et al., 2016, Migita et al., 2017)
FADS1 fatty acid desaturase 1	Unknown	
***HACD2* (*PTPLB*)** 3-hydroxyacyl-CoA dehydratase 2	Unknown	
***ACACA* (*ACSS1*)** acetyl-CoA carboxylase alpha	Tumour growth, malignancy (hepatocellular carcinoma)	(Tirinato et al., 2017)

(The gene changes related to positive outcomes are labelled with a yellow background; the gene changes which have no evidence to be associated with cancer are labelled with white background)

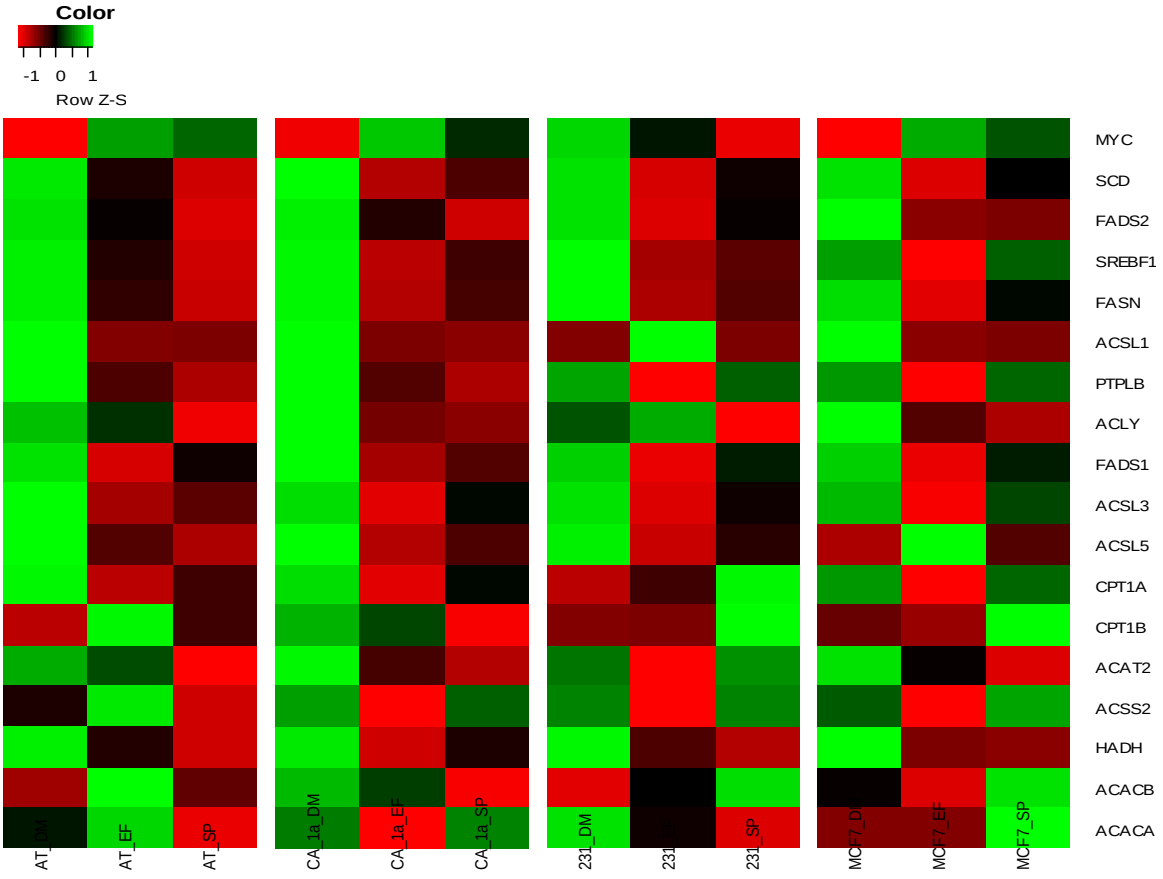


Figure 4-4 Heat-maps for presenting the expression levels of the key genes of fatty acid metabolism in different treatments and cancer cell lines
The key genes of fatty acid metabolism pathway are labelled on the right-hand side. And the colour code blocks present the relative expression levels of each gene in different conditions. The Z-score is to measure the relative number of the average number. The green colour represents higher RNA level; whereas, the red colour represents lower RNA level.

Fatty acid metabolism may play multiple roles in cancer; most of the crucial genes relevant to fatty acid metabolism have been linked to cancer cell proliferation. Several of these genes have been recognised as critical regulators of cancer stem cells (CSCs) and thus could be responsible for drug resistance, cancer invasion, metastasis, and cancer recurrence (Yi et al., 2018). *SCD* and *FASN*, which are associated with cancer survival and CSCs, are discussed below.

4.2.2.1.1 Stearoyl-CoA desaturase (*SCD*)

Among the candidate genes in fatty acid metabolism, *SCD* showed the most dramatic fold-change after antiretroviral drug treatment according to the ANOVA results. In TNBC cell lines, *SCD* was more than three times lower than the untreated control cells. Thus, it is worthwhile considering the possible role of *SCD* in antiretroviral drug-induced anti-cancer process.

There are several reasons why *SCD* could be a key factor linking antiretroviral drug treatment with anticancer processes. It has been shown to be one of the genes associated with various cancers, including breast cancer (Belkaid et al., 2015, Cui et al., 2017). It is also associated with breast cancer cell proliferation and tumour expansion in prostate and breast cancers (Kuo and Ann, 2018). Therefore, inhibiting *SCD* might potentially combat cancer. In this RNA-Seq experiment, both EFV- and SPV-treated cells were shown to significantly reduce *SCD* expression compared with the untreated control cells. This indicates that antiretroviral drugs can effectively suppress *SCD* expression, and it could explain how antiretroviral drugs eliminate TNBCs. Moreover, *SCD* is involved in CSC regulation (Tirinato et al., 2017). The overexpression of *SCD* has been linked to CSCs in many different cancers; whereas, decreasing *SCD* expression has been shown to affect CSC formation and to reduce their drug resistance. Drug resistance by CSCs is considered to be one of their major characteristics in lung and liver cancers (Yi et al., 2018). Therefore, the *SCD* suppression effect which was found in this RNA-Seq experiment is most likely relevant to CSC elimination.

4.2.2.1.2 Fatty Acid Synthase (*FASN*)

FASN is another fatty acid metabolism-related gene, which has been well-investigated in the past decade and is a confirmed therapeutic target for breast cancer (Menendez and Lupu, 2017a). No *FASN* inhibitor, however, has successfully passed clinical trials because of the drug efficiencies and the complexity of the human body (Flavin et al., 2010). *FASN* has also been implicated in HIV replication. Since HIV and LINE-1 share reverse transcription ability, it implies that *FASN* can be linked with LINE-1 regulation and thus play a role in antiretroviral drug-induced anti-TNBC effects. In this study, *FASN* expression was reduced by half after drug treatment in TNBCs suggesting that antiretroviral drugs might be able to regulate *FASN* expression directly or indirectly. However, more experiments need to be done to further explore this idea.

4.2.2.2 FoxO signalling pathway

The FoxO signalling pathway is considered important in antiretroviral drug treatment. Unlike fatty acid metabolism, the FoxO signalling pathway was only highlighted in the EFV down regulated gene set (eight genes, *q*-value = 0.0078) but not in the SPV down regulated gene set because of the selection criteria. The fold-change of catalase (*CAT*) in the SPV down regulated gene set was not as significant as in the EFV down regulated gene set. Cyclin-dependent kinase inhibitor 2D (*CDKN2D*) was increased rather than decreased after SPV treatment. However, serum and glucocorticoid regulated kinase 1 (*SGK1*), B-cell lymphoma 6 (*BCL6*), polo-like kinase 2 (*PLK2*), phosphoinositide-3-kinase, regulatory subunit 3 (*PIK3R3*), forkhead box O1 (*FOXO1*), and transforming growth factor beta 2 (*TGFB2*) passed the selection criteria (*q*-value less than 0.25, fold-change less than -1.2) in both the EFV and the SPV down regulated gene sets. The reduction of expression levels in *SGK1*, *BCL6*, and *PI3KR3*, which are associated with negative prognostic outcomes after antiretroviral drug treatment, implied a role for FoxO signalling pathway in the antiretroviral drug-induced anticancer process (Table 4-6). However, the decrease of *PLK2*, *TGFB2* and *FOXO1* after treatment led to conflicting results, since *PLK2* and *TGFB2* can act as either tumour suppressor or oncogene, depending on the stage and type of cancer (Dave et al., 2011, Ou et al., 2016). *FOXO1* has been considered both as a tumour suppressor gene and as an inducer of programmed cancer cell death (Lu and Huang, 2011, Farhan et al., 2017). The role of the FoxO signalling pathway in cancer can be very complicated. Many FoxO signalling pathway-related genes have been considered tumour suppressors; however, some new evidence suggests FoxO signalling pathway-related genes may be potential targets for treating cancer (Farhan et al., 2017, Hornsveld et al., 2018). The exact functions of the FoxO signalling pathway in cancer remain unclear.

Table 4-6 Antiretroviral drugs induced down regulated FoxO signalling pathway-related genes

Gene names	*Role in cancers*	*References*
EFV-treated or SPV-treated TNBCs		
SGK1 serum/glucocorticoid regulated kinase 1	Drug resistance (many cancers), tumour development (colorectal cancer)	(Talarico et al., 2016, Liang et al., 2017)
BCL6 B-cell lymphoma 6	Oncogene (leukaemia, lung and breast cancers)	(Cardenas et al., 2017, Wu et al., 2014)
PLK2 polo-like kinase 2	Tumour growth (colorectal cancer), oncogene (HCC), tumour suppressor (HCC)	(Ou et al., 2016)
PIK3R3 phosphoinositide-3-kinase, regulatory subunit 3 (gamma)	EMT inducer and metastasis promoter (colorectal, pancreatic cancer)	(Wang et al., 2014a, Peng et al., 2018)
FOXO1 forkhead box O1	Tumour suppressor	(Farhan et al., 2017, Hornsveld et al., 2018)
TGFB2 transforming growth factor, beta 2	Tumour suppressor in early stages; cancer promoter in advanced stages	(Dave et al., 2011)

(The gene changes related to positive outcomes are labelled with a yellow background; the gene changes related to negative outcomes are labelled with a grey-blue background; the gene changes related to either positive or negative outcomes are labelled with a red-orange background. HCC: hepatic cellular cancer.)

4.2.3 Up regulated pathways after antiretroviral drug treatment

After analysing down regulated genes and pathways involved in antiretroviral drug treatment, up regulated pathways potentially involved in antiretroviral drug treatment were also analysed. There were 262 genes (in appendix Table S4-3) that fitted the selection criteria (*q*-value less than 0.25, fold-change greater than 1.2) in the EFV up regulated gene set and 502 genes (in appendix Table S4-4) in the SPV up regulated gene set. Overall, up regulated pathways in both gene sets were more alike than the two down regulated gene sets. In EFV up regulated gene set, 246/262 genes had their protein products identified by the STRING-DB, and 231 potential links had been calculated among these proteins. The number of predicted links were greater than the number of expected links (168 links) and the protein-protein interaction (PPI) enrichment *p*-value was less than $2x10^{-6}$ suggesting that these proteins might at least partially have some biological connections. In SPV up regulated gene set, 460/502 genes had their protein products identified by the STRING-DB, and 630 potential links had been calculated among these proteins. The number of predicted links were greater than the number of expected links (457 links) and the protein-protein interaction (PPI) enrichment *p*-value was less than $4.95x10^{-11}$ suggesting that these proteins might have some biological connections. Genes associated with 'microRNAs in cancers' (EFV: eleven genes, *q*-value = 0.0004; SPV: eleven genes, *q*-value = 0.0447) (Figure 4-5, Table 4-7) and 'amino sugar and nucleotide sugar metabolism' (EFV: five genes, *q*-value = 0.0318; SPV: seven genes, *q*-value = 0.02) KEGG pathways were highlighted in both of the EFV up regulated and the SPV up regulated gene sets as shown in Figure 4-6 and Table 4-8.

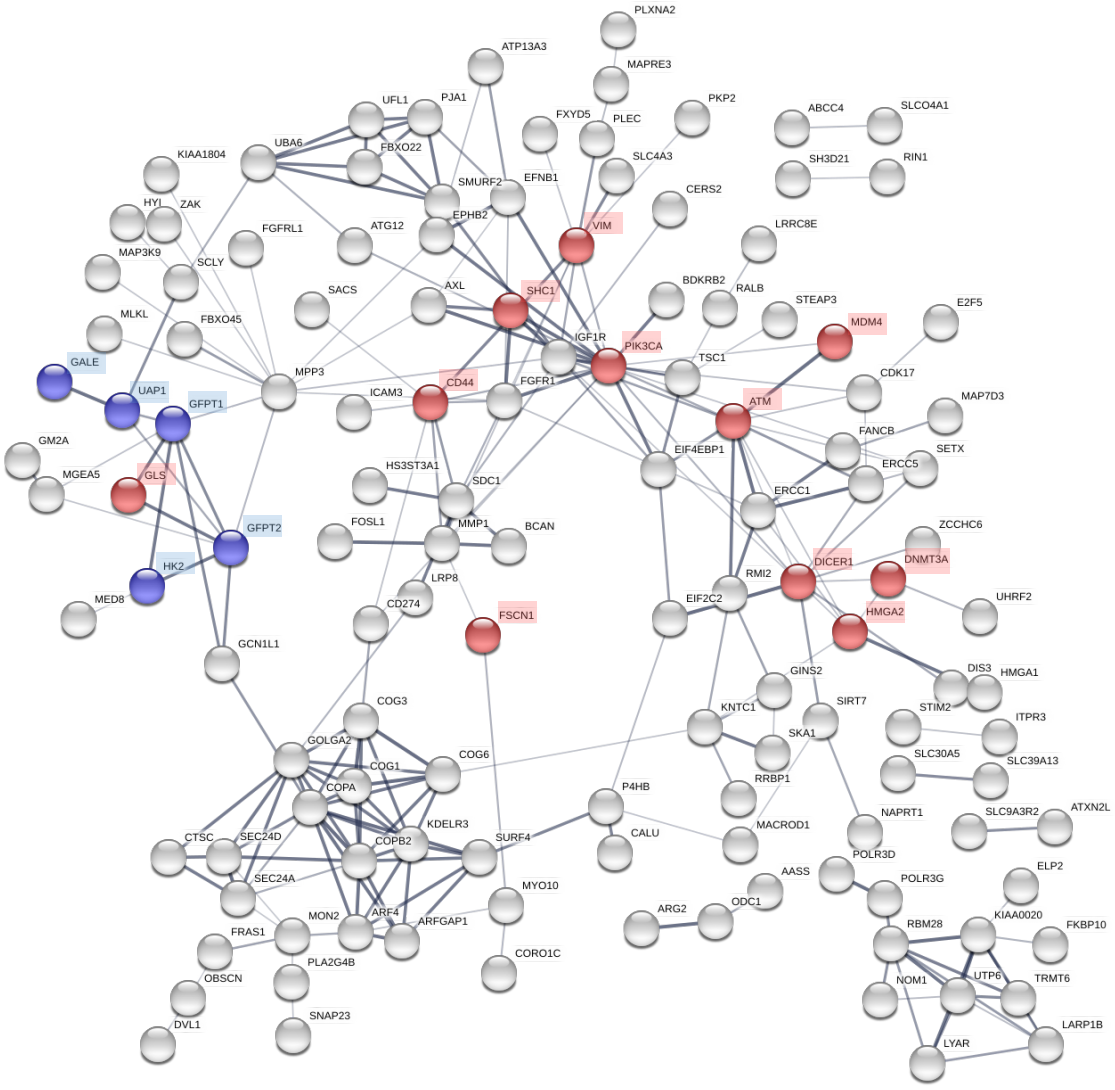


Figure 4-5 The STRING-DB gene network for EFV-treatment induced up regulated genes in TNBCs

This figure represents possible protein-protein interaction which was calculated by the STRING-DB software. A node represents a protein; a line represents a potential link. The thickness of the line indicates the strength of data support. The finest line represents medium confidence (0.400); the middle thickness line represents high confidence (0.700); the thickest line represents the highest confidence (0.900). The red nodes are proteins related to 'microRNAs in cancer'; the blue nodes are proteins related to 'Amino sugar and nucleotide sugar metabolism'.

Table 4-7 Highlighted KEGG pathways in EFV up regulated gene set by STRING-DB

Pathway ID	KEGG Pathway	Count in network	*q*-value
05206	MicroRNAs in cancer	11	0.0004
00520	Amino sugar and nucleotide sugar metabolism	5	0.0318

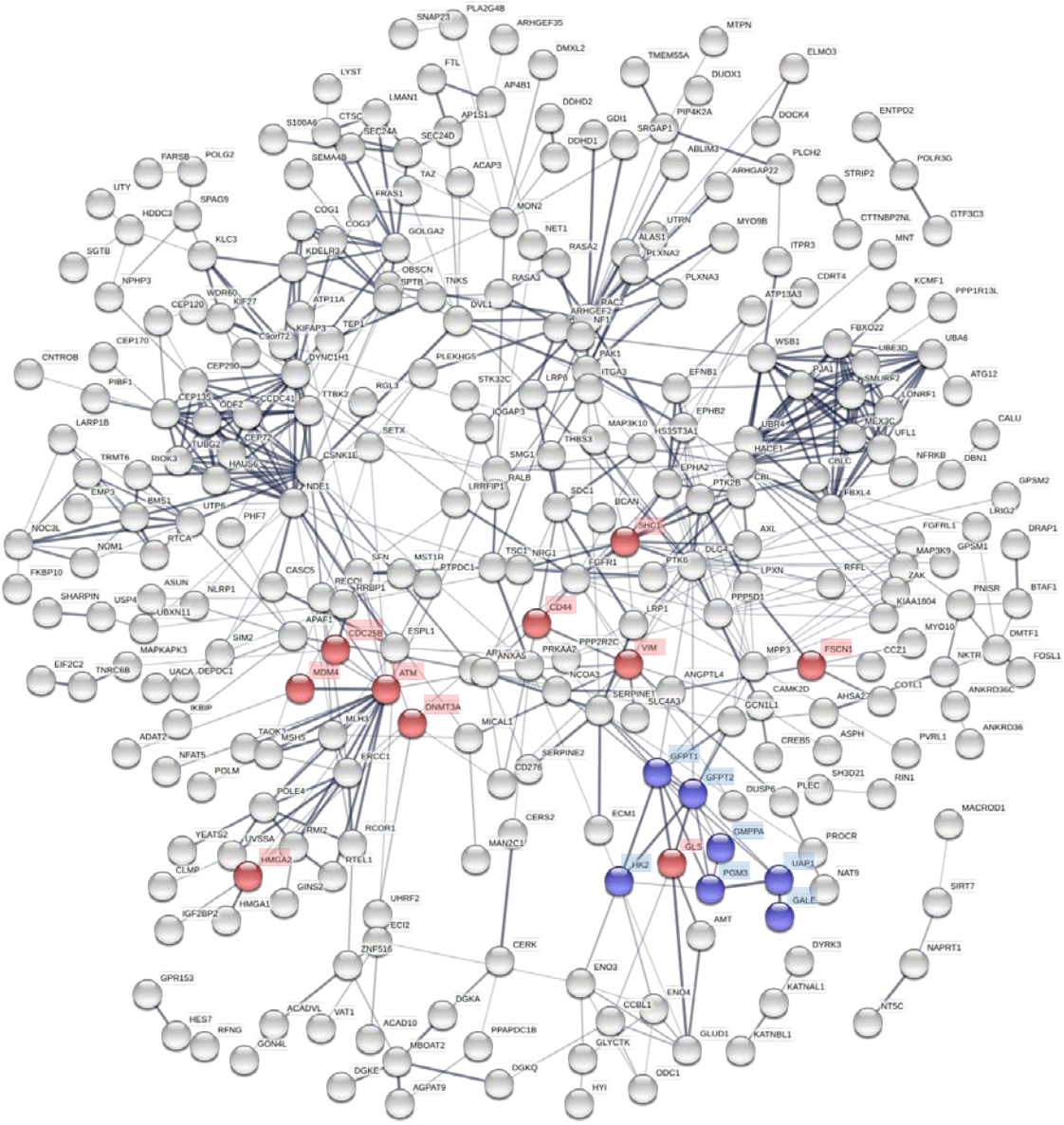


Figure 4-6 The STRING-DB gene network for SPV-treatment induced up regulated genes in TNBCs

This figure represents possible protein-protein interaction which was calculated by the STRING-DB software. A node represents a protein; a line represents a potential link. The thickness of the line indicates the strength of data support. The finest line represents medium confidence (0.400); the middle thickness line represents high confidence (0.700); the thickest line represents the highest confidence (0.900). The red nodes are proteins related to 'microRNAs in cancer'; the blue nodes are proteins related to 'Amino sugar and nucleotide sugar metabolism'.

Table 4-8 Highlighted KEGG pathways in SPV up regulated gene set by STRING-DB

Pathway ID	KEGG Pathway	Count in network	*q*-value
00520	Amino sugar and nucleotide sugar metabolism	7	0.02
05206	MicroRNAs in cancer	11	0.0447

4.2.3.1 Genes associated with cancer-related microRNAs

'MicroRNAs in cancer' represents a group of genes relevant to cancer-related microRNA regulations. The genes including high mobility group AT-hook 2 (*HMGA2*), vimentin (*VIM*), fascin homolog 1(*FSCN1*), cluster of differentiation 44 (*CD44*), glutaminase (*GLS*), mdm4 p53 binding protein homolog (*MDM4*), ataxia telangiectasia mutated (*ATM*), SHC-transforming protein 1 (*SHC1*), and DNA-methyltransferase 3 alpha (*DNMT3A*) were pointed out in both the EFV and the SPV up regulated gene set (Table 4-9). *HMGA2*, *VIM*, *FSCN1*, *CD44*, *GLS*, and *SHC1* have been connected to cancer growth and metastasis; while, *CD44* has also been recognised as one of the CSC markers in prostate, ovarian and breast cancers (Basakran, 2015). Interestingly, *DNMT3A* has been considered as an oncogene in hepatic cellular cancer (HCC) but also as a tumour suppressor gene in *acute myeloid leukaemia* (AML) (Chen and Chan, 2014, Brunetti et al., 2017). In contrast to the results observed for down regulating fatty acid metabolism-related genes, the up regulation of genes associated with cancer-related microRNAs is typically linked to poor prognostic outcomes.

Table 4-9 Antiretroviral drugs induced up regulated microRNAs in cancer-related genes

Gene names	***Role in cancer***	***References***
EFV or SPV-treated TNBCs		
HMGA2 high mobility group AT-hook 2	Cancer proliferation and metastasis, poor prognostic outcome	(Morishita et al., 2013, Wu et al., 2016b, Gao et al., 2017)
VIM vimentin	EMT marker (epithelial cancers)	(Satelli and Li, 2011)
FSCN1 fascin homolog 1, actin-bundling protein	Cancer development and metastasis (epithelial cancers)	(Hanker et al., 2013, Wang et al., 2017a)
CD44 a cluster of differentiation 44	Cancer stem cell marker, poor prognostic outcome (lung, breast and ovarian cancers)	(Basakran, 2015, Lin and Ding, 2017, Hu et al., 2018)
GLS glutaminase	Cancer growth	(Katt and Cerione, 2014, Lampa et al., 2017)
MDM4 Mdm4 p53 binding protein homolog	Negative tumour suppressor regulator	(Li and Lozano, 2013, Haupt et al., 2017, Bardot and Toledo, 2017)
ATM ataxia telangiectasia mutated	Cancer predisposition (lymphoid, gastric, breast cancers)	(Ahmed and Rahman, 2006, Choi et al., 2016)
SHC1 SHC-transforming protein 1	Cancer proliferation (breast and prostate cancers)	(Ravichandran, 2001)
DNMT3A DNA-methyltransferase 3 alpha	Oncogene (HCC), tumour suppressor gene (AML)	(Chen and Chan, 2014, Brunetti et al., 2017)

(The gene changes related to negative outcomes are labelled with a grey-blue background; the gene changes related to either positive or negative outcomes are labelled with a red-orange background. EMT: epithelial-to-mesenchymal transition. HCC: hepatic cellular cancer. ALM: adult acute myeloid leukaemia.)

4.2.3.2 The regulation of miR-21 and miR-182 after antiretroviral drug treatments

Since genes associated with 'microRNAs in cancers' were highlighted under antiretroviral drug treatment conditions, based on the STRING-DB search tool, it is worth investigating the relationship between microRNAs, antiretroviral drug treatment, and cancer malignancy.

Various microRNAs are involved in breast cancer regulation. According to the results from microRNA microarray experiments conducted in our laboratory (unpublished data from Dr Radhika Patnala, Dr Stephen Ohms, and Dr Danny Rangasamy), several microRNAs exhibited significant changes after EFV treatment (Figure 4-7). These experiments were based on a luminal-type breast cancer cell line T47D. The experimental design involved the treatment of T47D cells with EFV and then a comparison of variations in microRNA expression between EFV-treated and untreated T47D cells. They used microarray analysis to investigate the correlation between microRNAs and drug responses. Since 'microRNAs in cancer' had been shown in the up regulated gene set, it is worth examining whether the cancer-related microRNAs themselves can be up regulated in the drug-treated cells.

Two of the critical microRNAs showing vast changes after EFV treatment were microRNA 21 (*miR-21*) and microRNA 182 (*miR-182*). In the microRNA microarray assay, the EFV-treated T47D cells showed an increase in expression of 2.56 times for miR-21 and an increase of 2.58 times for *miR-182*, compared with the control cells (Figure 4-7). This suggested that *miR-21* and *miR-182* were strongly affected by EFV treatment. Thus, *miR-21* and *miR-182* were chosen for further investigation using microRNA quantitative reverse transcription polymerase chain reaction (qRT-PCR) in TNBCs.

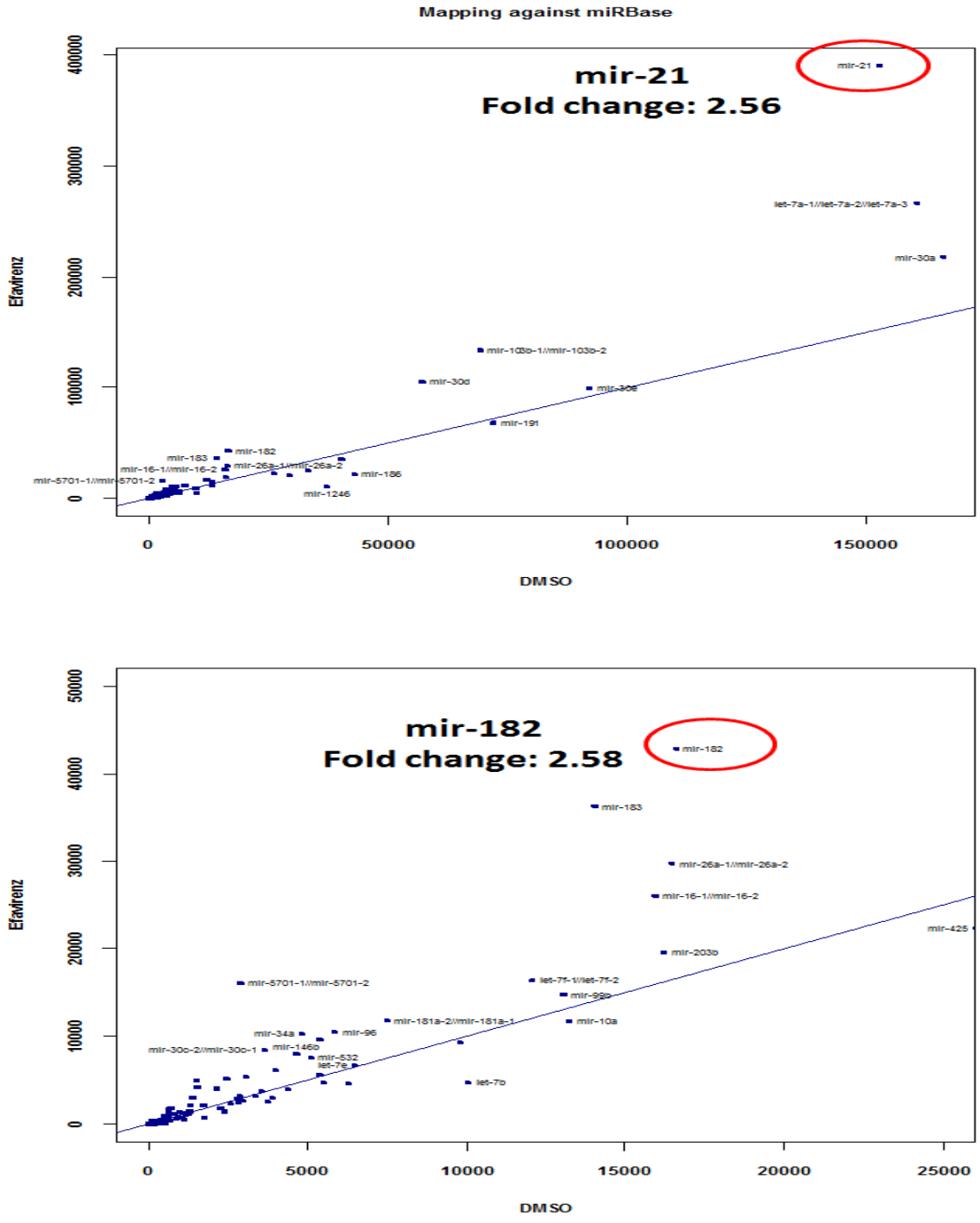


Figure 4-7 Scatterplots for gene expression between control and EFV treatment based on microarray data

The blue dots represent significant microRNAs detected in a microarray experiment. X-axis and y-axis represent the raw reads of the microRNAs. *miR-21* and *miR-182* (indicated by red circles) have greater fold-changes than other microRNAs.

Significant differences in *miR-21* and *miR-182* after EFV treatment were consistently detected in TNBCs. Both *miR-21* and *miR-182* increased after EFV treatment in MCF10AT, MCF10CA1α, and MDA-MB-231 cell lines (Figure 4-8). The following fold-change increases, in *miR-21*, of 0.61, 1.05, and 0.28 occurred in MCF10AT, MCF10CA1α, and MDA-MB-231 cells, respectively. There were fold-change increases, in *miR-182* of 0.97, 1.92, and 0.45 in MCF10AT, MCF10CA1α, and MDA-MB-231 cells, respectively. The increased expression of *miR-21* and *miR-182* in this experiment agreed with the findings of the microarray experiment in T47D cell lines.

miR-21 has been correlated with poor prognostic outcomes in breast cancers and is a biomarker of mesenchymal cancer cells (Yan et al., 2008, Pfeffer et al., 2015). It is also one of the CSC markers and may inhibit some tumour suppressor genes, including programmed cell death 4 protein (*PDCD4*), phosphatase and tensin homolog (*PTEN*) and others (Asangani et al., 2007, Zhang et al., 2010). Inhibiting miR-21 increases cell apoptosis and reduces cell proliferation (Pfeffer et al., 2015). Interestingly, our experiments showed that *miR-21* increased after antiretroviral drug treatment, either in the T47D microarray experiment or in the TNBC microRNA qRT-PCR experiment. This implied that the drugs may enhance the malignant potential of cells.

miR-182, however, displays dual roles in cancer regulation. Similar to *miR-21*, *miR-182* has been linked to cancer metastasis and unfavourable prognostic outcomes for cancer patients (Chiang et al., 2013, Lei et al., 2014). Yet, an alternative role for *miR-182* has recently been discovered: it can reduce certain drug resistance in breast cancers (Kouri et al., 2015, Li et al., 2018, Yue and Qin, 2019). Some microRNAs can present with multiple functions; acting as a tumour suppressor or as an oncogene in different circumstances (Ding et al., 2018). The exact functions of microRNAs in antiretroviral drug treatment are difficult to determine. Thus further experiments are essential before stating further conclusions.

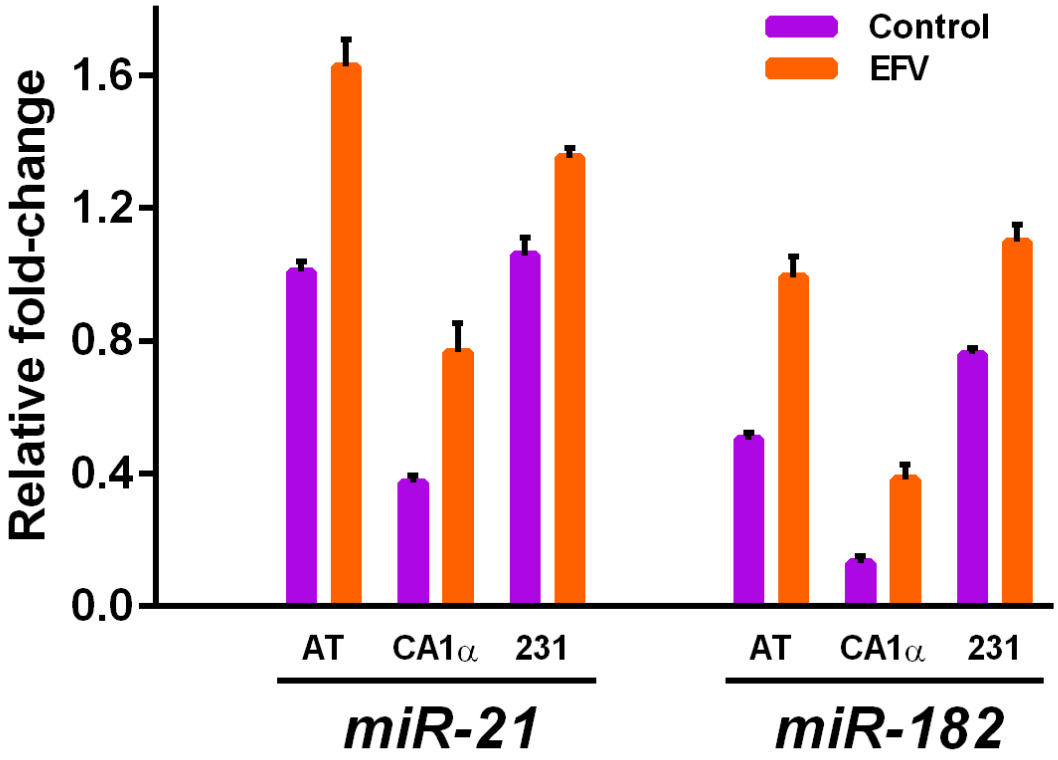


Figure 4-8 qRT-PCR data for comparison of *miR-21* and *miR-182* in different cell lines
miR-21 and *miR-182* are increased after drug treatment in all tested TNBC cell lines. The standard for the relative fold-change is the microRNA expression of untreated MCF10A cell line. Error bars: ±SD, n = 4.

4.2.3.3 Amino sugar and nucleotide sugar metabolism

Amino sugar and nucleotide sugar metabolism were also highlighted in up regulated gene sets (Table 4-10). Several of the candidate genes listed in the amino sugar and nucleotide sugar metabolism pathway, including UDP-N-acetylglucosamine pyrophosphorylase 1 (*UAP1*), hexokinase 2 (*HK2*), and glutamine-fructose-6-phosphate transaminase 1 (*GFPT1*), were associated with poor prognostic outcomes. *GFPT1* can also be considered as a suppressor of epithelial-to-mesenchymal (EMT) transition and metastasis in gastric cancers (Duan et al., 2016). *HK2* which is a key gene for cancer initiation and growth is the most well-known among these candidate genes. It has been linked to glycolytic regulation, which could be associated with cancer growth (Wang et al., 2016, DeWaal et al., 2018). There is evidence suggesting that *HK2* inhibition or depletion could be an effective strategy in treating many different cancers (Wu et al., 2017). Few reports show the relationship with *UAP1* and cancer; however, a current study (Itkonen et al., 2014) suggests that *UAP1* may be overexpressed in prostate cancers and protect cancer cells from cell stress. It might be improving cancer growth under cancer treatment. To date, little is known about any direct link between amino sugar and nucleotide sugar metabolism and cancers.

Table 4-10 Antiretroviral drugs induced up regulated amino sugar and nucleotide sugar metabolism-related genes

Gene names	***Role in cancers***	***References***
EFV or SPV-treated TNBCs		
UAP1 UDP-N-acetylglucosamine pyrophosphorylase 1	Cancer growth, drug resistance	(Itkonen et al., 2014)
GFPT2 glutamine-fructose-6-phosphate transaminase 2	Unknown	
HK2 hexokinase 2	Tumour initiation and growth	(Patra et al., 2013, Wang et al., 2016, DeWaal et al., 2018, Wu et al., 2017)
GFPT1 glutamine-fructose-6-phosphate transaminase 1	Poor prognosis marker (Pancreatic cancer, HCC), EMT and metastasis suppressor (Gastric cancer)	(Yang et al., 2016, Duan et al., 2016, Li et al., 2017)
GALE UDP-galactose-4-epimerase	Unknown	

(The gene changes related to negative outcomes are labelled with a grey-blue background; the gene changes related to either positive or negative outcomes are labelled with a red-orange background; the gene changes which have no evidence to be associated with cancer are labelled with white background. EMT: epithelial-to-mesenchymal transition. HCC: hepatic cellular cancer)

4.2.4 The role of p53 in antiretroviral drug treatment

The combination of the down regulated and the up regulated gene sets may also identify critical genes pathways which were not identified in the individual down regulated or up regulated gene set. While some genes are up regulated, others are down regulated in a given pathway; hence if we focus on a specific regulating side, we may miss out on the big picture. Although by enlarging the gene pool we may decrease the sensitivity of the STRING-DB software and reduce the significance of pathways, it is important to analyse the combined gene sets for EFV and SPV treatments in order to discover the masked genes and pathways.

Our analysis combined the significant genes (*q*-value less than 0.25) with fold-changes greater than 1.2 and less than -1.2 (EFV: 499 genes; SPV: 881 genes), and found that most of the pathways being highlighted overlapped with the down regulated or the up regulated gene sets (Figure 4-9, Table 4-11, and in appendix Figure S4-1, Table S4-5). 'Fatty acid metabolism' was highlighted in both EFV- and SPV-treated cells; while, 'FoxO signalling pathway' and 'microRNAs in cancers' was only indicated in EFV-treated cells. The genes involved in these pathways were very similar to the genes described in individual gene sets. However, Tumor protein p53 *(TP53)* appeared in the significant 'microRNAs in cancer' pathway when down regulated and up regulated gene sets were combined; but not in any of the individual gene sets. This meant that unlike other genes associated with the 'microRNAs in cancer' pathway, *TP53* was down regulated after antiretroviral drug treatment.

TP53 is the most intensely studied tumour suppressor gene and has multiple functions in cancer. *TP53* can act as a negative regulator of proliferation in cancer cells via apoptosis, and it can also be an EMT and a cell cycle regulator (Varna et al., 2011). It plays roles in ageing, autophagy, mitotic catastrophe, and angiogenesis (Varna et al., 2011). *TP53* has long been recognised as an oncogene (Soussi and Wiman, 2015); however, scientists now acknowledge that the oncogenic functions of p53 protein are caused by mutant p53 protein rather than wild-type p53 protein. The mutant protein has lost its tumour suppressor abilities and induces cancer progression (Muller and Vousden, 2014). *TP53* mutation has been found to exist in more than half of human cancers (Parrales and Iwakuma, 2015). Frequently, *TP53* has a missense mutation enabling the stable expression of mutant p53 protein within tumour cells. In numerous situations, these mutant p53 proteins are unable to target p53-responsive genes thus terminating

oncogene suppression in cancers (Liu, 2011). Other functions of mutant p53 protein have also been confirmed in cancers which explains why p53 mutation is a crucial area of cancer research (Freed-Pastor and Prives, 2012). Thirty per cent of breast cancers overall and sixty per cent of TNBCs carry the mutated *TP53* gene (Bertheau et al., 2013, Turner et al., 2013). The mutant p53 protein is related to cancer development and has been associated with therapeutic resistance to chemotherapy, hormonotherapy and radiotherapy (Varna et al., 2011). It has been strongly linked with unfavourable prognostic outcomes in TNBCs and may be responsible for cancer initiation (Varna et al., 2011). Therefore, the targeting of mutant p53 is considered to be a potential therapeutic strategy.

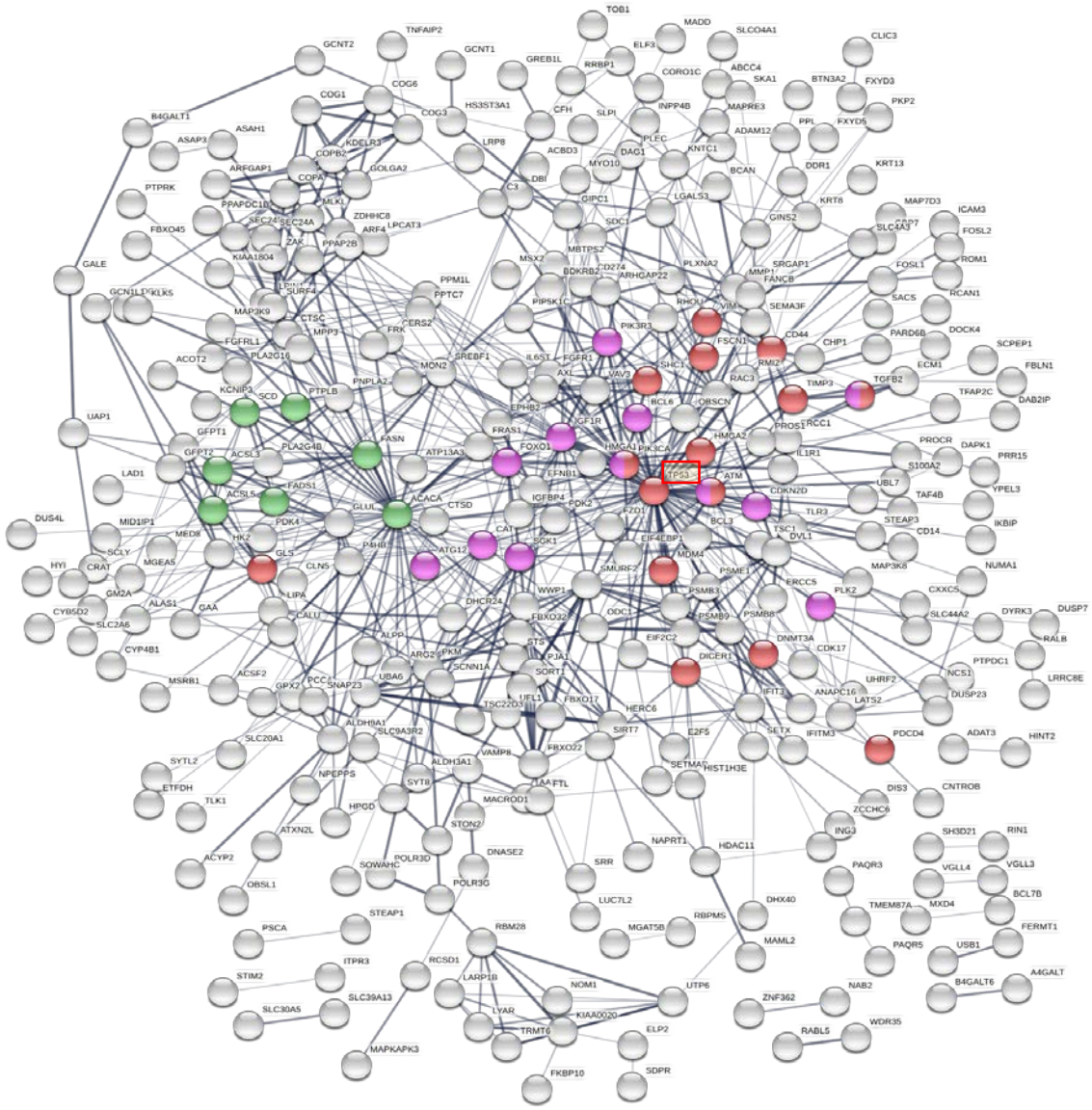


Figure 4-9 The STRING-DB gene network for EFV-treatment induced up regulated and down regulated genes in TNBCs

This figure represents possible protein-protein interaction which was calculated by the STRING-DB software. A node represents a protein; a line represents a potential link. The thickness of the line indicates the strength of data support. The finest line represents medium confidence (0.400); the middle thickness line represents high confidence (0.700); the thickest line represents the highest confidence (0.900). The red nodes are proteins related to 'microRNAs in cancer'; the pink nodes are proteins related to 'FoxO signalling pathway'; the green nodes are proteins related to 'fatty acid metabolism'.

Table 4-11 Highlighted KEGG pathways in EFV down regulated and up regulated gene sets using STRING-DB

Pathway ID	Pathway description	Count network	*q*-value
05206	**Micro RNAs in cancer**	15	0.000376
04068	**FoxO signalling pathway**	12	0.00349
01212	**Fatty acid metabolism**	7	0.011

Antiretroviral drug treatment specifically reduced the expression of *TP53* in three tested TNBC cell lines. In this RNA-Seq study, it had been predicted that antiretroviral drug treatment may universally repress the expression of *TP53* according to the significant reduction of *TP53* in ANOVA analysis. However, contrary to the expectation, *TP53* was only suppressed in MCF10AT, MCF10CA1α, and MDA-MB-231 cell lines after the drug treatment. Antiretroviral drug-treated TNBC cells expressed approximately two-third of *TP53* RNA compared with their untreated-controls; whereas, only subtle differences in *TP53* RNA levels were detected in MCF10A and MCF7 cell lines (Figure 4-10). This result indicates that antiretroviral drugs might have very limited effects on *TP53* RNA in MCF10A and MCF7 cell lines which contain wild-type *TP53* gene (Lim et al., 2009). Unlike MCF10A and MCF7 cell lines, the MDA-MB-231 cell line expresses the mutant *TP53* gene (Lim et al., 2009), whereas no report has been published on whether MCF10AT and MCF10CA1α express the wild-type or mutant *TP53* gene. These observations imply that either antiretroviral drugs can suppress *TP53* RNA expression specifically in TNBCs, or that MCF10AT and MCF10CA1α cells contain mutant *TP53*, as do MDA-MB-231 cells, that can be inhibited by antiretroviral drugs.

Based on our RNA-Seq results and reading of the published literature, we have proposed the following hypothesis to explain the diverse morphological changes observed after treatment of the TNBC cell lines with antiretroviral drugs. Antiretroviral drugs may specifically suppress the expression of both mutant and wild-type *TP53* genes in TNBCs. The MDA-MB-231 cells contain mutant *TP53* and express mutant p53 protein (Lacroix et al., 2006); whereas, the MCF10AT and the MCF10CA1α cells, similar to MCF10A cells, contain wild-type *TP53* and express wild-type p53 protein (Lim et al., 2015). Mutant p53 protein is well-recognised as a potential therapeutic target for cancer (Parrales and Iwakuma, 2015), and suppressing it in MDA-MB-231 cells might effectively reduce cancer proliferation. In contrast, knocking-down wild-type p53 may induce EMT (Jiang et al., 2011, Wang et al., 2013). When MCF10AT and MCF10CA1α cells were treated with the drugs, they displayed mesenchymal markers such as SLUG and fibronectin (mentioned in chapter 3) which may be linked to poor prognostic outcomes. This suggests that extensive suppression in TNBCs of p53 may be responsible for the altered morphological changes observed in antiretroviral drug-treated TNBC cells.

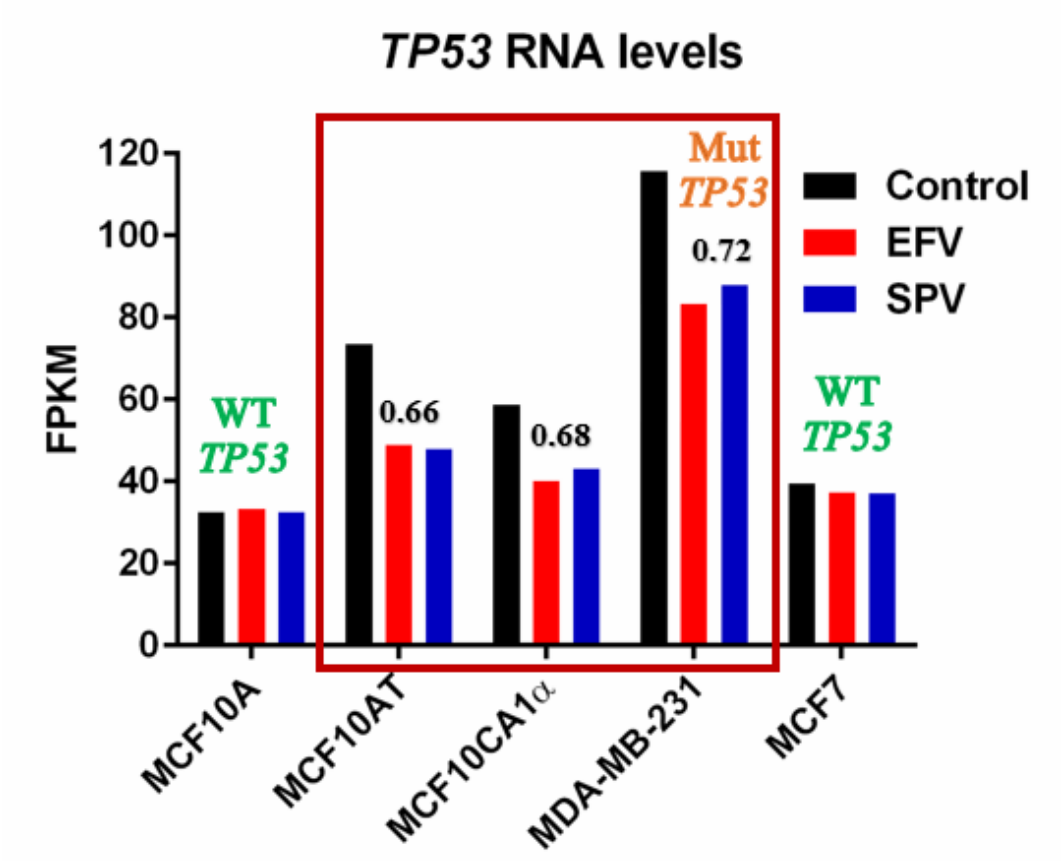


Figure 4-10 The RNA-Seq result of *TP53* RNA expression in different cell lines
The y-axis represents the reads of Fragments per Kilobase Million (FPKM). The black-bars represent untreated-controls; the red-bars represent EFV treatments; the blue-bars represent SPV treatments. The antiretroviral drug-treated cells express around 70% of *TP53* RNA compared with untreated-control cells in MCF10AT, MCF10CA1α, and MDA-MD-231 cell lines. MCF10A and MCF7 express wild-type *TP53* RNA; whereas MDA-MB-231 expresses mutant *TP53* RNA. There is no evidence to suggest which type of *TP53* RNA is expressed by MCF10AT and MCF10CA1α.

Interestingly, EFV, LINE-1, and p53 can be connected by several studies. p53 has been shown to influence the early replication of HIV by inhibiting HIV RT (Shi et al., 2018). This RT shares a similar function to that of LINE-1 RT. In addition, the p53 binding site has been identified in the LINE-1 promoter region (Harris et al., 2009). Therefore, p53 might be relevant to LINE-1 regulation. Recent studies suggested that p53 co-evolved with endogenous transposons, including LINE-1, in the germline millions of years ago. p53 regulated their function in the genome, possibly by a positive-feedback loop (Levine et al., 2016). Therefore, it might be possible that the expression of LINE-1 affects the expression level of p53 by a feedback signal. Moreover, EFV has been linked with p53-induced autophagy-related cell death in human keratinocyte cells, (Dong et al., 2013), which links EFV treatment and p53 regulation. Taken together, these findings suggest that EFV treatment can affect both LINE-1 and p53, while LINE-1 and p53 may interact with each other. The exact relationship and the mechanisms involved are still vague, and more evidence and experiments are needed to test our hypothesis. Although this topic is beyond the scope of this project, it could be a further direction for researchers who want to understand the role of p53 in antiretroviral drug-induced cancer activities.

4.2.5 Other important genes in antiretroviral drug induced cancer responses

Although the importance of fatty acid metabolism and genes associated with cancer-related microRNAs have been illustrated, many other genes may have a role to play in how cancer cells respond to antiretroviral drugs. When analysing the gene sets via STRING-DB, we found that some genes were highlighted in various groups according to the GO database. In contrast to the KEGG database, the GO database focuses more on the locations and functions of proteins rather than their pathways. Based on the GO database, many genes with closely-related functions which might further explain the phenotypes described in the previous chapter were pointed out and are reviewed at the end of this chapter. Furthermore, as mentioned, network analysis also provided some important information, the results of which were considered. Finally, genes displaying highly statistical significances in ANOVA analysis but which were not involved in any particular pathways and networks were listed and discussed.

4.2.5.1 Cell morphology-related genes

As shown in the previous chapter, genes related to morphological changes in the cells were emphasised in two up regulated gene sets. Many genes related to cell morphology were highlighted in antiretroviral drug-treated up regulated gene sets suggesting that antiretroviral drugs might be able to induce morphological changes in cancer cells. This result was observed with immunofluorescence and western blotting and discussed in the previous chapter. Consistent with previous observations, increases were observed not only in cell projection-related genes, but also in cell migration-related genes, after antiretroviral drug treatment. Cell projection is a sign of cell differentiation and these results suggested that the cells tested tended to differentiate after antiretroviral drug treatment. This is consistent with previous findings that some antiretroviral drugs might be able to transit cancer cells from mesenchymal states to epithelial states, thus reducing the malignancy of the cancers. Cell migration has been linked to cancer metastasis and the expression of the cell migration genes may lead to unfavourable prognostic outcomes. Many modules in our network analysis also displayed similar trends with focal adhesion, cell junction, and stress fibre highlighted in three out of eleven TNBC-only modules (in appendix Figure S4-2). The RNA-Seq data was consistent with the phenotypes observed in the previous chapter, in which some TNBCs, especially MCF10AT and MCF10CA1α, showed signs of favourable and unfavourable prognostic outcomes after antiretroviral drugs treatment.

4.2.5.2 The role of mitochondria in antiretroviral drug-induced cancer responses

The network analysis of the TNBC+MCF7 gene set extended the analysis in an alternative direction. There were eight modules that passed the selection criteria in the TNBC+MCF7 gene set (in appendix Figure S4-3), and only three of them indicated significant protein-protein association networks (in appendix Table S4-6). In agreement with the network analysis for the TNBC-only gene set, the modules based on the TNBC+MCF7 gene set identified many genes related to cell morphological changes, including adherens junction, cell projection, and focal adhesion genes (Figure S4-4). There were many genes which corresponded mitochondria in some modules (in appendix Table S4-7).

This result suggested the potential involvement of mitochondria in antiretroviral drug-induced cancer responses. EFV seems to affect mitochondrial complex I (Lee and Boelsterli, 2014, Apostolova et al., 2014) which might further affect the balance between oxidised and reduced nicotinamide adenine dinucleotide (NAD^+/NADH) (Lee and Boelsterli, 2014). This NAD^+/NADH balance is crucial for CSC maintenance and thus influences the malignancy of cancers (Bonuccelli et al., 2017). Additionally, EFV has been shown to damage the mitochondrial complex I (Purnell and Fox, 2014). Thus, a possible link between EFV and mitochondrial complex I regulation has been illustrated. Moreover, LINE-1 suppression affects a piRNA which might target human mitochondrial Hsp 70 (Mortalin) and thereby regulating the functions of mitochondria (Ohms and Rangasamy, 2014). Mitochondria might also play a role in antiretroviral drug-induced cancer responses; however, this is beyond the scope of this project and more studies are needed to support this hypothesis.

4.2.5.3 CSC-related genes

The importance of CSC was drawn to our attention when several CSC-related genes were highlighted in RNA-Seq analysis. Interestingly, among those genes which changed significantly with treatment, CSC-related genes such as *CD44*, *MED8*, *DMXL2*, and *PROCR*, as well as miR-21, increased, while others such as *CHMP4B*, *ACSL3*, *ALDH9A1*, *SCD*, and *FASN* decreased. The inconsistent results emphasised the complexity of CSCs within TNBCs. Therefore, a further study of CSC was conducted, and it is described in the next chapter.

Table 4-12 CSC-related genes involved in antiretroviral drug-induced cancer responses

Gene names	*Drug response*	*Role in cancers*	*Reference*
CD44 a cluster of differentiation 44	↑	CSC indicator (lung, breast and ovarian cancers)	(Basakran, 2015, Lin and Ding, 2017, Hu et al., 2018)
MED8 mediator complex subunit 8	↑	Downstream of Notch signalling pathway (CSC regulator), metastasis (renal cell, lung cancers)	(Dewi et al., 2011)
DMXL2 Dmx-like protein 2	↑	EMT and CSC indicator	(Alison et al., 2010, Faronato et al., 2015)
PROCR protein C receptor	↑	EMT and CSC indicator	(Hwang-Verslues et al., 2009)
CHMP4B charged multivesicular body protein 4B	↓	drug resistance (hepatocellular cancer)	(Hu et al., 2015a)
ACSL3 acyl-CoA synthetase long-chain family member 3	↓	promoting cancer initiation (lung cancer)	(Padanad et al., 2016)
ALDH9A1 aldehyde dehydrogenase 9 family member A1	↓	CSC indicator (renal cancer)	(Henrion et al., 2015, Chang et al., 2018)
SCD stearoyl-CoA desaturase	↓	CSC indicator (epithelial cancer)	(Tirinato et al., 2017)
FASN fatty acid synthase	↓	CSC indicator (epithelial cancer)	(Tirinato et al., 2017)

(The gene changes related to positive outcomes are labelled with a yellow background; the gene changes related to negative outcomes are labelled with a grey-blue background. EMT: epithelial-to-mesenchymal transition. CSC: cancer stem cell.)

4.3 Discussion

In the previous chapter, some contentious results were analysed and discussed. Our major experiments suggested that antiretroviral drug-treated cancer cells become more normal; whereas, other data including a part of immunofluorescence and western blotting results suggested that antiretroviral drug-treated cancer cells could be linked to unfavourable prognostic outcomes such as cancer metastasis, drug resistance and cancer recurrence. In an effort to understand the conflicting results, RNA-Seq experiments were undertaken to investigate the changes induced in the RNA by antiretroviral drugs.

In this chapter, RNA-Seq analysis detected quite a few possible regulatory pathways in breast cancer cell lines which may be influenced by antiretroviral drug treatment. We compared the variation in the entire genome's RNA expression for different TNBC cell lines when treated with antiretroviral drugs and when untreated. According to the STRING-DB analysis, 'fatty acid metabolism', 'FoxO signalling pathway', 'microRNAs in cancer', and 'amino sugar and nucleotide sugar metabolism' were involved in antiretroviral drug-induced cancer responses. Also, the network analysis indicated that the cell morphology regulating genes were strongly altered during the drug treatment. Moreover, we hypothesised that there is a link between antiretroviral drug-induced mitochondrial damage and anticancer effects.

One significant pathway being highlighted in both EFV and SPV treatments was 'fatty acid metabolism'. Fatty acid metabolism has been proven to play a key function in cancer (Tirinato et al., 2017): it is linked to cancer metastasis, therapeutic resistance, and cancer stemness (Kuo and Ann, 2018). A demonstration version of the STRING-DB, STRING-DB version 11.0 (https://string-db.org/), was released very recently. This new version takes into account the changes of every imported gene in order to enhance its prediction function. It also adds two other databases as references, the Reactome pathway database (https://reactome.org/) and the UniProt Knowledgebase (https://www.uniprot.org/), thereby further improving its analysis. According to this latest STRING-DB, 'fatty acid metabolism' was the only pathway being highlighted in EFV treatment gene set. It had arisen not only in the KEGG pathway but also in the Reactome and the UniProt pathways. Similarly, 'fatty acid metabolism had been identified in the SPV treatment gene set. Although the results from the Reactome and the UniProt databases suggested that SPV might not only influence fatty acid metabolism and also affect

interferon, ‘fatty acid metabolism’ was the only pathway being highlighted in the KEGG pathway in the STRING-DB version 11.0. These results strongly implied that fatty acid metabolism is the key pathway involved in antiretroviral drug-induced anti-TNBC effects. Our analysis presented a vast majority of fatty acid metabolism-related gene inhibition as a result of antiretroviral drug treatment, which is in agreement with the published literature (Belkaid et al., 2015, Menendez and Lupu, 2017a). This indicated that the anti-breast cancer effects of antiretroviral drugs could be associated with fatty acid metabolism dysfunction. This result not only highlights a further direction for TNBC studies but also points to a possible way to improve anti-FASN drugs.

The experiment of examining FASN expression in drug-treated TNBCs had been considered. Because of the time constraints and a shortage of materials, we decided to focus on cancer stem cell (CSC) regulation. However, it is worth to examine the regulations of the fatty acid metabolism pathway in EFV-treated TNBCs in the future.

Pathways linking to unfavourable prognostic outcomes were also identified by RNA-Seq analysis. Increasing expression of several cancer- and CSC-related genes was detected in antiretroviral drug treated-cancer cells. These results suggested that RNA-Seq data opened up more complex questions, and therefore a huge gap in our knowledge needs to be filled. Before we can conduct further investigations, we must consider some limitations of the methods used.

The major limitation of ANOVA analysis is that the variation of every sample set might be ignored despite different cell lines having different responses to drugs because of their genetically diverse backgrounds. By using ANOVA analysis, a complete picture is observed but it lacks data from a single cell line. It indicates that significant genes (with the highest fold-change) could mask information from individual cell lines. If a gene in an individual cell line shows no response to a particular drug but exhibits strong responses in other cell lines, this gene would be recognised as significant.

There are also some limitations in the network analysis which should be considered. The major problem is that some important information for individual genes may be lost when considering ‘significant change’. The other limitation of network analysis is that gene responses may not be consistent even in the same pathway. For instance, in a particular pathway, some genes are up

regulated while others are down regulated, and this resulted in them being put into different modules. If fewer genes were placed into the same module, the potential link within modules may be difficult to interpret. This may be why the morphologically related genes are the only ones observed in the TNBC-only network analysis. Because the changes in cell morphology are dominant, other regulatory pathways might be masked. This may indicate that many relevant genes are altered to induce cellular phenotypes. If more data were collected, more tolerance could be built into the network analysis and minor differences accommodated.

RNA-Seq analysis is a preliminary step for assessing the possible pathways involved. More experiments need to be performed in order to confirm the results. Thus, in order to validate these results an attempt to generate a LINE-1 overexpression cell line was made to examine whether LINE-1 inhibition was the cause of fatty acid metabolism down regulation. The MCF10A non-cancer cell line only expresses limited amounts of *SCD*, *FASN* and LINE-1. If an overexpressing LINE-1 MCF10A cell line is generated, it could test whether *SCD* and *FASN* are regulated by LINE-1. However, because endogenous LINE-1suppression mechanisms exist in the MCF10A cell line, it is very difficult to successfully generate this cell line. Additional methods need to be considered to overcome the challenge.

Regardless of all these limitations, the importance of heterogeneity within TNBCs is deliberated in this study. According to the RNA-Seq results, antiretroviral drugs may potentially induce an anticancer process by inhibiting fatty acid metabolism in breast cancers. However, some other pathways involved can be linked to poor prognostic outcomes. A number of CSC-related genes have been highlighted in RNA-Seq analysis, and some of these increase after antiretroviral drug treatment. This raises the following questions: did antiretroviral drugs kill most of the non-CSCs, so that CSCs became the majority in the surviving population? or did antiretroviral drugs actually induce the development of CSCs?

To answer these questions, comparisons between the responses of the CSC population and the entire population was considered; the details of which are outlined in the next chapter.

Chapter 5 The effects of EFV and SPV on breast cancer stem cells

5.1 Introduction

Cancer stem cells (CSCs) are a small group of cancer cells that contain stem cell-like characteristics. These include an ability for self-renewal and differentiation into different types of cancer cells. More importantly, they are capable of initiating tumours (Chen et al., 2013). Researchers are paying increasing attention to CSCs in cancers because they are thought to play a significant role in cancer recurrence and metastasis (Clevers, 2011). CSCs may also be responsible for tumours developing resistance to traditional cancer therapeutic methods, such as chemotherapy and radiotherapy (Wu and Alman, 2008, McDermott and Wicha, 2010, Bielecka et al., 2017). Thus, specifically targeting CSCs could be a potential strategy for treating and preventing cancers.

The results presented in the previous chapters clearly emphasised the importance of CSCs in antiretroviral drug-induced cancer responses. The evidence from drug response experiments suggested a critical role for cancer heterogeneity and implied that a small population of cancer cells may respond differently to the drugs. Furthermore, the RNA-Seq experiments showed that CSC-related genes could be altered upon treatment with antiretroviral drugs.

In Chapter 3, differentiation phenotypes, which usually identify normal cells, were observed in antiretroviral drug-treated MCF10AT, MCF10CA1α, MDA-MB-231, and T47D cells. In addition, antiretroviral drug-treated MCF10AT and MCF10CA1α cells showed mesenchymal-cell phenotypes which are usually correlated with malignant cancers. This finding was supported by immunofluorescence and western blot data showing that EFV-treated TNBCs exhibited some mesenchymal cell characteristics, such as increasing fibronectin and SLUG proteins, decreasing CD24 protein, and the relocation of E-cadherin protein. These data were in contrast to the original prediction that antiretroviral drugs might be able to reverse cancer cells to normal cell-like status (Sbardella et al., 2011). Although some of the TNBC cells did return to being non- cancerous, others expressed obvious cancer phenotypes. These results, therefore, emphasised the complexities arising from cancer heterogeneity in TNBC cells and suggested that some cells might become more cancerous after antiretroviral drug treatment.

In Chapter 4, RNA-Seq data further confirmed that TNBCs express epithelial non-CSC markers, mesenchymal markers and CSC markers simultaneously after the drug treatment. The expression of epithelial markers successfully represents antagonising cancer growth. Mesenchymal and CSC markers are associated with poor prognostic outcomes and can be linked to cancer recurrence and cancer metastasis. These data agreed with the immunofluorescence and western blot data which indicated the possibility of increased levels of breast CSCs (BCSCs) in the whole cancer population after the treatment. This then gave rise to the question of whether antiretroviral drugs eliminated most of the non-CSCs leaving only highly drug-resistant CSCs to survive, or whether the drugs may induce CSCs by some unknown mechanisms. To address these questions, it is important to compare the drug responses of the whole cancer cell population, the CSC-only population, and the non-CSC population.

In this chapter, the effects of antiretroviral drug treatment on the CSC population were investigated for several TNBC cell lines, including MCF10AT, MCF10CA1α, MDA-MB-231, BT-549, and BT-20. Because CSC classification methods can vary, initially the CSC population was sorted from the whole cancer population using the two most well-known CSC identification methods, Aldehyde Dehydrogenase (ALDH) activity and CD44/CD24 characteristics. The percentage of CSCs in the whole population was determined for both antiretroviral drug-treated cells and untreated cells. The method of CSC isolation is addressed and the disadvantages of CSC isolation and the difficulties for CSC maintenance is explained. Finally, CSC tumorsphere functional assays are discussed and the responses of CSCs to the drugs investigated. The chapter concludes with a discussion of the effects of antiretroviral drugs in BCSCs.

5.2 Results

5.2.1 Antiretroviral drug responses of CSCs in triple-negative breast cancer cell lines

Over the past few decades, several CSC identification methods have been discovered; however, the lack of universal CSC markers has hindered CSC research. The main reason for this is that CSC markers vary depending on the type of cancer and also on the subtype of certain cancers (Wu and Alman, 2008). Moreover, CSCs identified by different methods may exhibit various phenotypes, enrich alternative genes, and have minimal overlap with each other (Liu et al., 2014a). It is therefore critical to evaluate and compare diverse clusters of CSCs when undertaking CSC research.

The most commonly used markers for determining BCSC are CD44 and CD24 cell surface antigens. The breast cancer cells with high CD44 and low CD24 expression have stem cell-like characteristics and are recognised as CSCs (Sheridan et al., 2006). Other types of BCSCs with different features are known as '*mesenchymal-like* CSCs' (Liu et al., 2014a). Additionally, breast cancer cells with high aldehyde dehydrogenase (ALDH) activities have been named '*epithelial-like* CSCs' (Liu et al., 2014a). Furthermore, cells divergent from the main population in fluorescence-activated cell sorting (FACS) analysis have been defined as 'side population'. This subset of cells exhibits a higher chemical efflux capacity, compared to the main population of cancers, which may contribute to their drug resistance (Wu and Alman, 2008). Recently, the CSC tumorsphere functional assay has become a popular tool in the field of CSC research. There is evidence to suggest that such three-dimensional culture conditions are relatively similar to *in vivo* microenvironments and that they might simulate CSC growth under pseudo-physiological conditions.

5.2.1.1 Epithelial-like CSCs

Changes in the *epithelial-like* CSC population, after antiretroviral drug treatment, were examined in this study. Significant suppression of *ALDH9A1* RNA, a member of ALDH, in antiretroviral drug-treated TNBCs was observed in RNA-Seq analysis. The ALDEFLUOR kit (Stemcell Technology) was utilised to detect cells with higher aldehyde dehydrogenase activities, which is the *epithelial-like* CSCs (Figure 5-1). The proportion of *epithelial-like* CSCs was reduced after EFV treatment in all tested breast cancer cell lines, yet very little decrease was observed in the non-cancerous MCF10A cell line (Figure 5-2). The percentage of *epithelial-like* CSCs decreased from 7.56±0.71% to 4.53±0.79% in MCF10AT cells; from 66.30±8.97% to 20.76±7.51% in MCF10CA1α cells; from 11.06±1.96% to 3.87±1.05% in MDA-MB-231 cells; and from 16.76±1.17% to 2.32±0.86% in T47D cells (Table 5-1). The *epithelial-like* CSCs were reduced by more than 1.5-fold, after EFV treatment, in all tested cancer cell lines. Notably, the *epithelial-like* CSCs in EFV-treated MCF10CA1α and MDA-MB-231 cells were decreased 3-fold compared to their untreated controls. As predicted, T47D, a non-TNBC control, showed a 7-fold decrease in its *epithelial-like* CSC population after EFV treatment. This implied that the *epithelial-like* CSCs in luminal breast cancer, particularly T47D cell, might be more sensitive to the *epithelial-like* CSCs in TNBCs. All these cancer cell lines display significant differences when exposed to untreated- and EFV-treated conditions. The *p* values of the Student's *t*-test (paired, two-tails) were 0.033, 0.018, 0.038, and 0.006 for MCF10AT, MCF10CA1α, MDA-MB-231, and T47D, respectively (Table 5-1). In the non-cancerous control, MCF10A, the percentage of the $ALDH^{high}$ CSCs was decreased from 4.08±1.17% to 3.13±1.50%. This difference was not statistically significant (*p*-value = 0.233) for EFV-treated and untreated cells (Figure 5-2 and Table 5-1). These data demonstrate that the proportion of the *epithelial-like* CSCs can be reduced by EFV treatment.

Consistent with the results of EFV treatment, the SPV-treated breast cancer cells included a relatively small $ALDH^{high}$ CSC population compared with the untreated controls (in appendix Figure S5-1). After SPV treatment, the percentage of the $ALDH^{high}$ CSCs was reduced: by 3.6% in MCF10AT; by 17.9% in MCF10CA1α; by 2.0% in MDA-MB-231; and by 10.6% in T47D cancer cell lines. In contrast, there was only 0.1% difference between SPV-treated and untreated-MCF10A cells. However, because of the shortage of the SPV compound, no replicate was performed and therefore no statistical analysis was available.

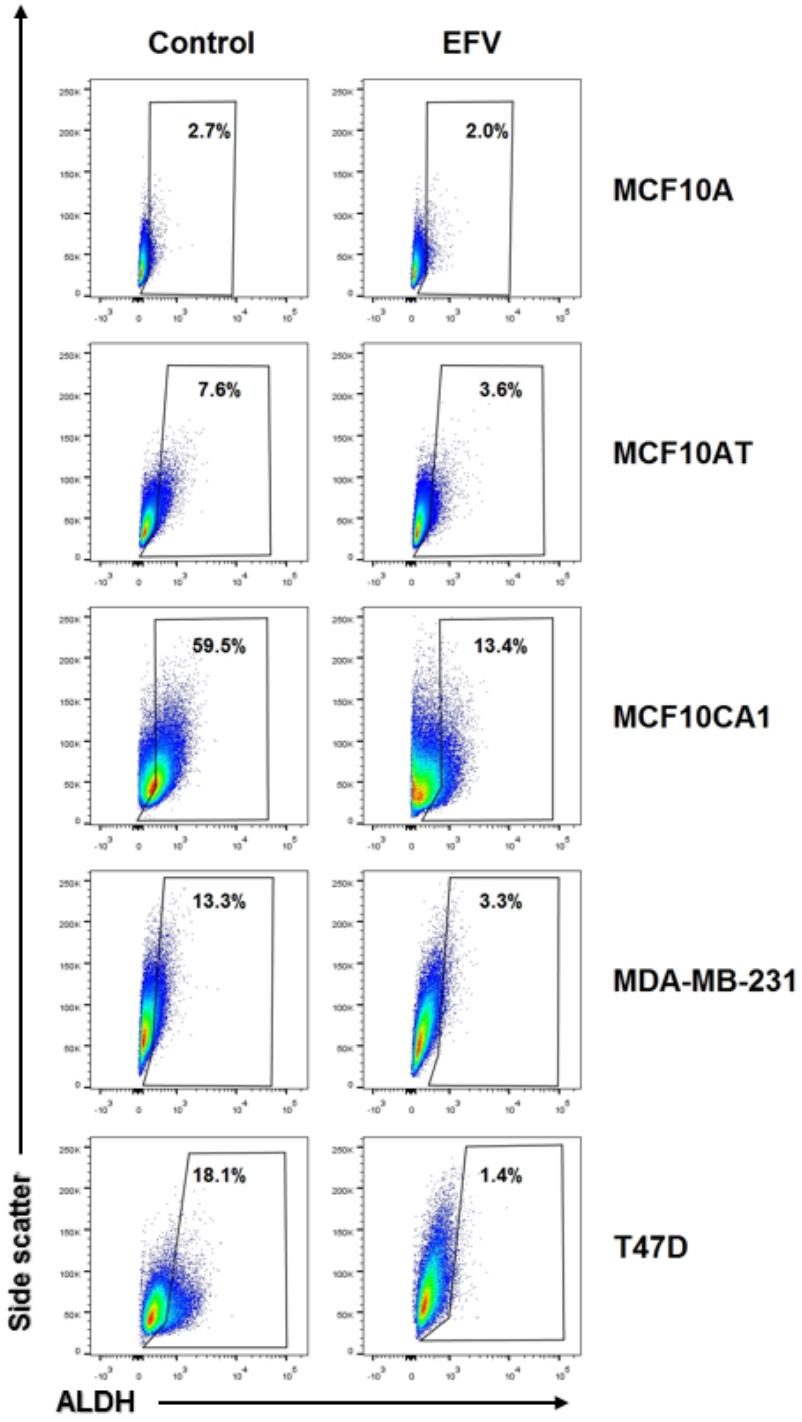


Figure 5–1 The ALDH^{high} CSCs detected using flow cytometry in EFV-treated and untreated-MCF10A, MCF10AT, MCF10CA1α, MDA-MB-231, and T47D cells

The dots inside the frame lines represent the cells with high ALDH activities. The frame lines were set based on their ALDH suppression controls.

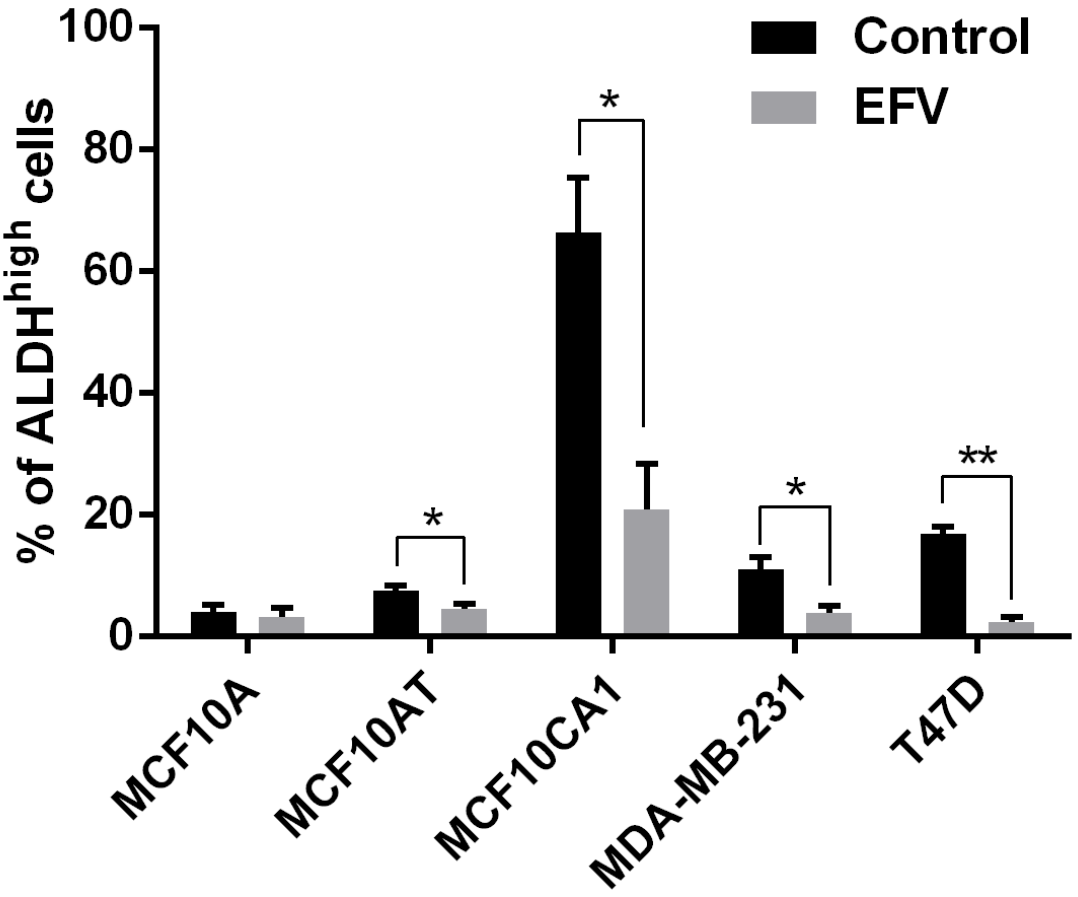


Figure 5–2 The percentages of the *epithelial-like* CSCs presented in EFV-treated and untreated cancer cells

This figure is a quantitative bar chart for the ALDEFLUOR™ results in TNBC cell lines. The black bars represent untreated control cells; the grey bars represent EFV-treated cells. Significant differences were detected in all EFV-treated cancer cell lines. Two-tailed, paired Student's *t*-test was conducted for statistical analysis. **p*<0.05; ***p*<0.01. Error bars: ±SD, n = 3.

Table 5-1 The percentages of ALDH^{high} cells presented in EFV-treated and untreated population

		% of ALDH^{high}			AVERAGE	SD	*p*-value
MCF10A	Control	4.56	4.93	2.74	4.08	1.17	0.23261
	EFV	2.55	4.83	2.00	3.13	1.50	
MCF10AT	Control	7.61	6.82	8.24	7.56	0.71	0.03259*
	EFV	3.64	4.79	5.15	4.53	0.79	
MCF10CA1α	Control	62.89	76.47	59.53	66.30	8.97	0.01823*
	EFV	28.43	20.44	13.42	20.76	7.51	
MDA-MB-231	Control	13.27	10.4	9.51	11.06	1.96	0.03773*
	EFV	3.26	5.09	3.27	3.87	1.05	
T47D	Control	16.12	16.05	18.11	16.76	1.17	0.00619**
	EFV	3.14	2.40	1.42	2.32	0.86	

(Two-tail, paired Student's *t*-test, **p*<0.05; ***p*<0.01)

These results showed that EFV and SPV could be effective drugs for reducing the *epithelial-like* CSCs in both TNBC and non-TNBC cell lines. However, increasing expression of CSC-related genes such as *CD44*, *HK2*, *DMXL2*, *and PROCR* could not be explained by the *epithelial-like* CSC results. If the *epithelial-like* CSC represents all CSCs within the whole cancer population, the CSC-related genes should largely be down-regulated, unless they are being activated by alterative molecular pathways unrelated to CSC regulation. Therefore, other types of CSCs were investigated to better understand the whole picture of CSC responses to the drugs.

5.2.1.2 Mesenchymal-like CSCs

The *mesenchymal-like* CSCs were investigated in order to observe the changes resulting from the drug treatment. The results of CD44/CD24 staining suggested an increase in the *mesenchymal-like* CSC population upon treatment with antiretroviral drugs. In this study, the percentage of the $CD44^+/CD24^-$ cells was increased after EFV treatment in all tested TNBC cell lines (Figure 5-3, 5-4). This trend was opposite to the result of the *epithelial-like* CSCs. In MCF10AT cells, the $CD44^+/CD24^-$ population increased from 3.69±1.14% to 9.81±1.64%; in MCF10CA1α, the $CD44^+/CD24^-$ population increased from 29.49±0.57% to 45.83±4.69%; and in MDA-MB-231, the $CD44^+/CD24^-$ population increased from 87.64±0.78% to 92.35±1.79% (Table 5-2). A 1.1-fold to 2.7-fold increase in the *mesenchymal-like* CSC population in TNBC cell lines was observed. Very few *mesenchymal-like* CSCs were detected in MCF10A (non-cancerous control) and T47D (non-TNBC control). Statistical analysis of these results showed significant differences in the population of *mesenchymal-like* CSCs under EFV-treated and untreated-conditions for MCF10AT, MCF10CA1α, and MDA-MB-231 cells. The *p*-values of the Student's *t*-test (two-tailed, paired) were 0.004, 0.002, and 0.004 for MCF10AT, MCF10CA1α, and MDA-MB-231, respectively (Table 5-2). A similar trend was observed when TNBC cells were treated with SPV (in appendix Figure S5-2); however statistical analysis could not be performed owing to the shortage of SPV. These data indicated the $CD44^+/CD24^-$ mesenchymal-CSC populations were altered by EFV or SPV drug treatments, and that different cell lines respond differently to the drugs.

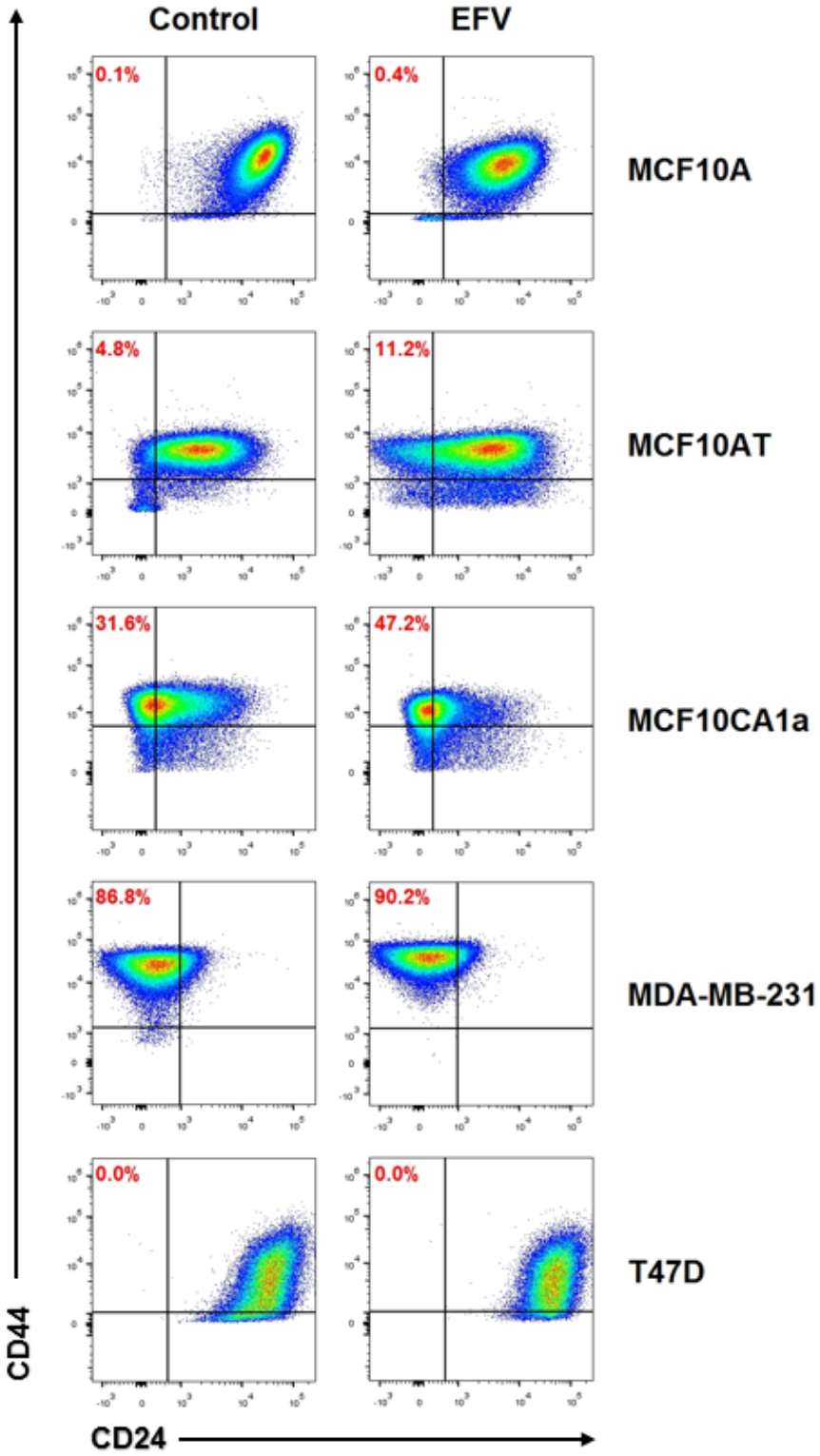


Figure 5–3 The CD44⁺/CD24⁻ cells detected in untreated- and EFV-treated MCF10A, MCF10AT, MCF10CA1α, MDA-MB-231, and T47D cells

The dots in the up-left corner represent the CD44⁺/CD24⁻ *mesenchymal-like* CSCs. The frame lines were set based on their unstained controls.

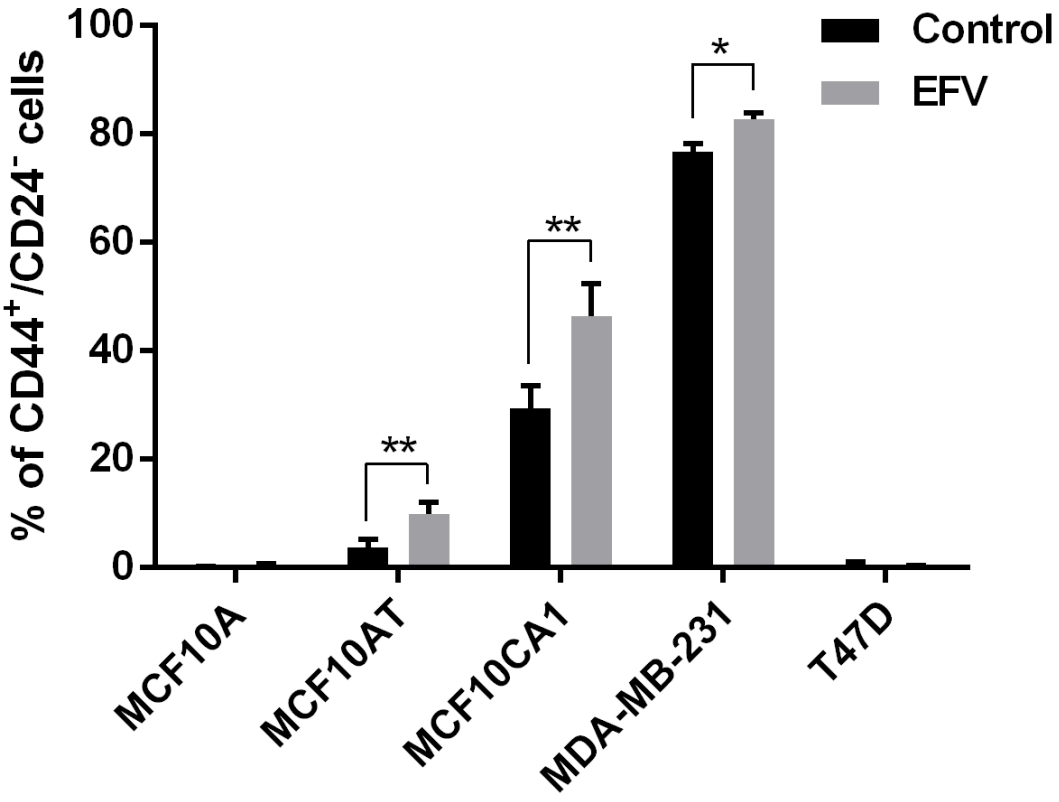


Figure 5–4 The percentages of $CD44^+/CD24^-$ cells presented in untreated and EFV-treated cells
This figure is a quantitative bar chart for the CD44/CD24 staining results in TNBC cell lines. The black bars represent untreated control cells; the grey bars represent EFV-treated cells. The significant differences were detected in all EFV-treated TNBC cell lines. Two-tailed, paired Student's *t*-test is conducted for statistical analysis. *p<0.05; **p<0.01. Error bars: ±SD, n = 3.

Table 5-2 The percentages of $CD44^+/CD24^-$ cells presented in EFV-treated and untreated cells

		% of $CD44^+/CD24^-$			AVERAGE	SD	*p*-value
MCF10A	control	0.01	0.08	0.06	0.05	0.03	0.21371
	EFV	0.02	0.44	0.75	0.40	0.30	
MCF10AT	control	2.11	4.21	4.75	3.69	1.14	0.00360**
	EFV	7.50	10.78	11.15	9.81	1.64	
MCF10CA1α	control	31.58	24.47	32.43	29.49	3.57	0.00190**
	EFV	47.21	39.52	50.76	45.83	4.69	
MDA-MB-231	control	88.68	87.46	86.79	87.64	0.78	0.02281*
	EFV	94.59	92.26	90.20	92.35	1.79	
T47D	control	0.00	0.99	0.31	0.43	0.41	0.30985
	EFV	0.00	0.28	0.15	0.14	0.11	

(Two-tailed, paired Student's *t*-test, *p<0.05; **p<0.01)

5.2.1.3 Side population

As previously mentioned, CSCs are made up of a diverse set of population cells, one of which is the 'side population' (SP). The SP describes a minor population of cells which has stem cell-like characteristics including self-renewal and an ability to survive after chemotherapy (Wu and Alman, 2008). These types of cells express large amounts of multidrug-resistance-type ATP-binding cassette protein (ABC) transporters. They are, therefore, able to efflux more Hoechst 33342 nucleotide staining dye than other cells (Wu and Alman, 2008, Bielecka et al., 2017). It is possible that SP cells are more resistant to chemotherapy than the *epithelial-like* and the *mesenchymal-like* CSCs because of their drug efflux capabilities (Nakanishi et al., 2010).

There are several reasons why it was very difficult to identify SP cells in our tested cell lines: The SP is an extremely small population in TNBC and T47D cell lines (less than 1% in total population) compared with most of the non-TNBC cell lines (Nakanishi et al., 2010). In addition, the insufficient staining and gating techniques, as well as dynamic dye efflux process, may greatly affect the results (Wu and Alman, 2008, Golebiewska et al., 2011). Furthermore, the cytotoxicity of the Hoechst 33342 SP staining dye, i.e., its mutagenicity and the damage to DNA that it causes, may also be responsible for the difficulties experienced in studying these cells (Wu and Alman, 2008). Moreover, EFV- or SPV- treated cells were unhealthy, compared to the untreated cancer cells. As a consequence and not surprisingly, very few SP cells were observed in all cell lines tested using the typical SP identification method (in appendix Figure S5-3). These difficulties make the stem-like 'side population' unfeasible to study in triple-negative CSCs.

5.2.2 Exploring *mesenchymal-like* CSC results

Given that the *epithelial-like* CSC and the *mesenchymal-like* CSC data gave rise to conflicting results, a critical examination of the methodology of CSC determination was required. One consideration was that the cell surface proteins including CD44 and CD24 were compromised by cell apoptosis and necrosis. Moreover, it is possible that drug treatment influenced the CD44/CD24 staining and the accuracy of the results. It is also possible that antiretroviral drugs may increase the CSC population through mechanisms unrelated to LINE-1 inhibition. The following experiments were conducted to address these questions.

5.2.2.1 Alteration of functional CSCs after drug treatment

The tumorsphere formation assay was employed to determine whether the increasing percentage of *mesenchymal-like* CSCs in drug-treated cancer cells was directly linked to the destruction of the cell surface membrane. Only the cells with self-renewal ability can survive and proliferate in low-nutrition MammoCult Medium. These cells were postulated to be functional CSCs and were strongly correlated with the *mesenchymal-like* CSCs in certain breast cancer cell lines (Wang et al., 2017b). The experiment involved treating cells with DMSO (untreated-control), EFV, and SPV for four days, after which the healthy cells were harvested (in appendix Figure S5-4). This was followed by seeding the same number of cells ($1x10^4$) in MammoCult medium and culturing the cells for seven days to form tumorspheres. The tumorspheres were then photographed, counted and their size measured for each individual treatment.

The results (Figure 5-5 and Table 5-3) suggested that EFV- and SPV-treated MCF10AT and MCF10CA1α cells formed more tumorspheres and their size was larger than their controls. The SPV-treated MCF10CA1α cells, in particular, produced more than twice the number of tumorspheres than untreated cells. Although the tumorsphere-forming cells might not strictly correlate with *mesenchymal-like* CSCs, this data can still partially confirm that the treatment can increase the CSC numbers in the entire cancer population. Therefore, the integrity of the cell membrane might be less relevant to the observation that the *mesenchymal-like* CSC increase after the drug treatment.

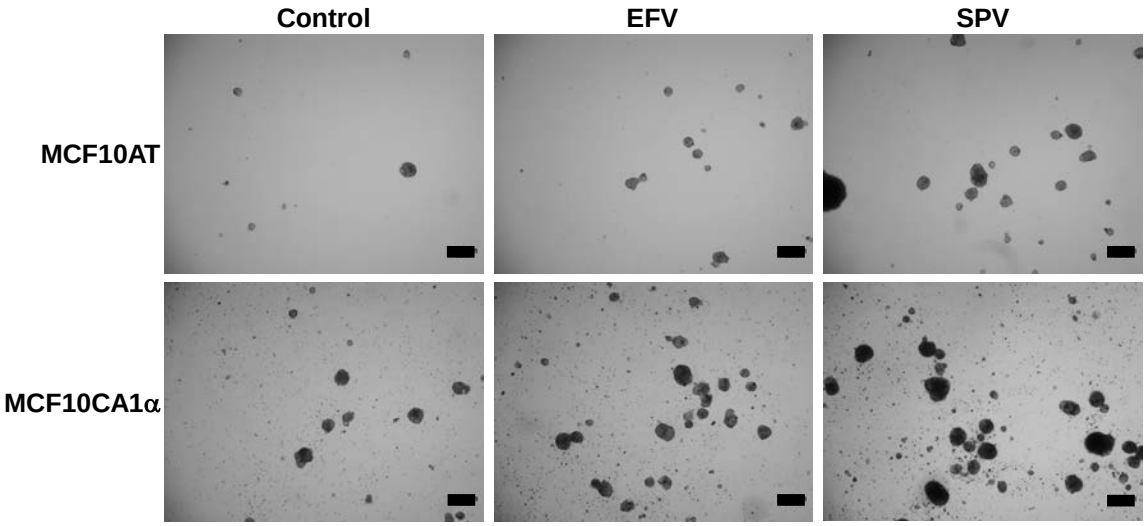


Figure 5–5 Tumorsphere formation after antiretroviral drug treatment

MCF10AT and MCF10CA1α cells were originally cultured in traditional two dimensional-culture and treated with antiretroviral drug. After the treating period, the remaining cells were harvested and seeded into MammoCult™ three-dimensional culture. Under this culture condition, the remaining cells grew after a few days. At the seven-day time point, the entire culture plate was photographed by Olympus IX81 inverted brightfield/fluorescence microscope. Scale bar: 200 µm

Table 5-3 The numbers of tumorspheres formed from different pre-treatment conditions

Number of spheres /10^4 cells seeded		60-100 µm	100-200 µm	200-400 µm	>400 µm	total
MCF10AT	**Control**	49	37	6	0	92
	EFV	49	42	8	0	99
	SPV	75	38	12	1	126
MCF10CA1α	**Control**	92	133	32	0	257
	EFV	177	189	57	0	423
	SPV	229	213	81	2	525

5.2.2.2 Mesenchymal-like CSCs and LINE-1 siRNA inhibition

LINE-1 short hairpin RNA (shRNA) inhibition was employed to exclude any potential side effects caused by the anti-retroviral drugs. A pUTR plasmid, encoding an shRNA sequence targeting the LINE-1 promoter, was transfected into MCF10AT and MCF10CA1α cells (as described in Chapter 2) in order to inhibit LINE-1 expression in the cells. The non-functional empty vector, pSM2 plasmid, was transfected as a control for the pUTR plasmid. By comparing the drug inhibition with the direct shRNA inhibition of LINE-1, it was hoped to better understand the activity of the antiretroviral drugs in CSC enrichment.

Even though the siRNA inhibition results were not straightforward to interpret due to the extremely fast cell proliferation rate of the cancer cells, there was an increase in the *mesenchymal-like* CSC population upon transfection with the pUTR plasmid (Figure 5-6). The percentage of the *mesenchymal-like* CSCs was 0.5% in MCF10AT-pSM2 cells and 9.3% in MCF10AT-pUTR cells; whereas, it was 14.5% in MCF10CA1α-pSM2 and 27.8% in MCF10CA1α-pUTR cells. Both MCF10AT and MCF10CA1α with partial LINE-1 silencing by pUTR presented with much smaller numbers of *mesenchymal-like* CSCs compared to their controls. Therefore, LINE-1 inhibition through shRNA seemed to increase the *mesenchymal-like* CSC population in MCF10AT and MCF10CA1α cell lines which were consistent with the results arising from drug treatment. Hence, the increase in the *mesenchymal-like* CSCs in TNBC cell lines, induced by EFV- or SPV-treatment, was attributed to LINE-1 inhibition.

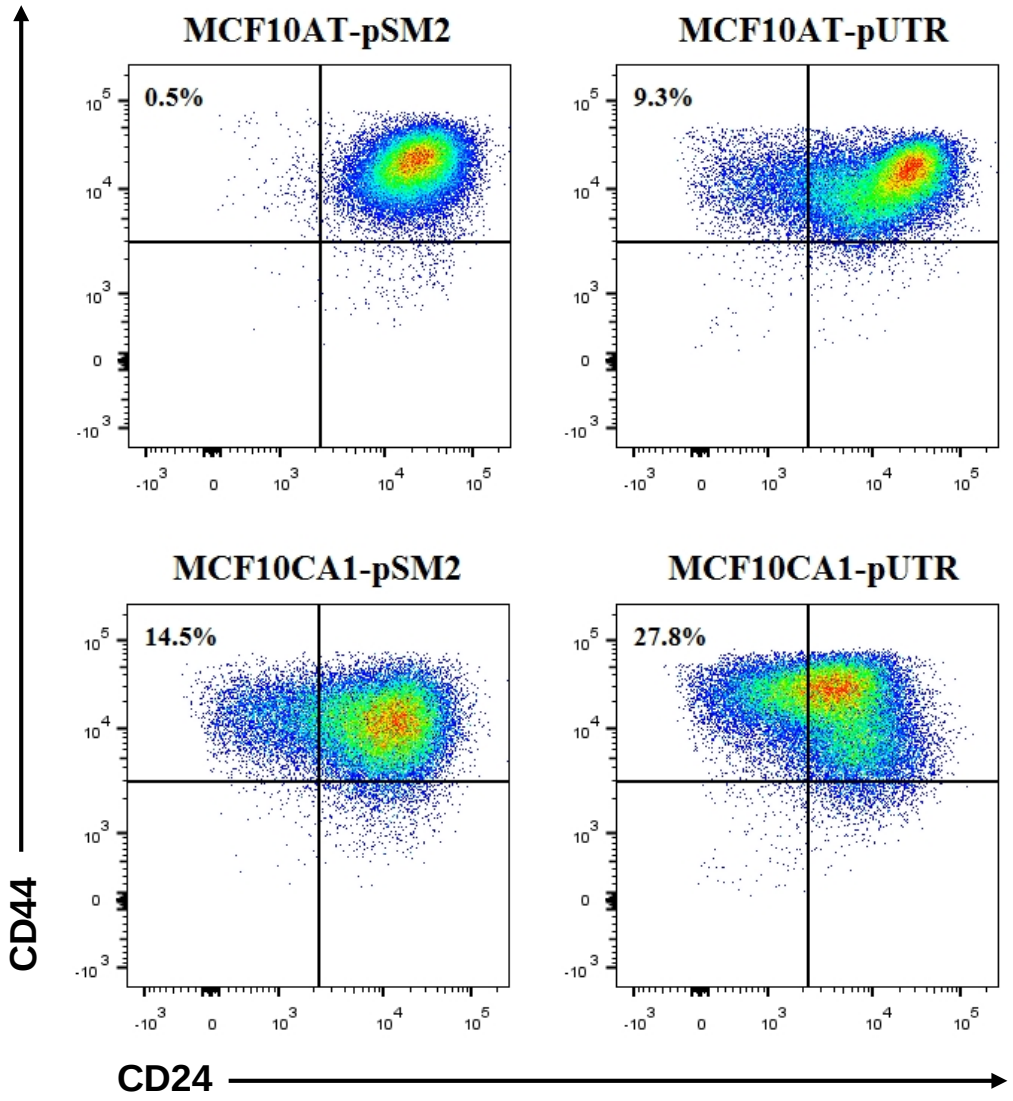


Figure 5–6 The *mesenchymal-like* CSCs in MCF10AT-pSM2, MCF10AT-pUTR, MCF10CA1α-pSM2, MCF10CA1α-pUTR cells

The dots in the up-left corner of each FACS figure represent the $CD44^{+}/CD24^{-}$ *mesenchymal-like* CSCs. The frame lines were set based on their unstained controls.

5.2.3 Cancer Stem Cell isolation

To further understand the basis for the contradictory results between the *epithelial-like* and the *mesenchymal-like* CSC populations after the drug treatment (see section 5.2.1.1 and section 5.2.1.2), a pure CSC population was sorted. Thus, CSC isolation may be an appropriate strategy to provide some clarity on the CSC population.

5.2.3.1 Isolating $CD44^+/CD24^-$ mesenchymal-like CSCs

The $CD44^+/CD24^-$ and the $CD44^-/CD24^-$ cancer cells were isolated and collected by a cell sorter. The intracellular staining methods, such as ALDEFLUOR assay, may be more harmful to the cells than the cell surface marker staining methods, CD44 and CD24 staining. Furthermore, these unhealthy sorted cells can be difficult to re-culture. The observation of an increasing population of the $CD44^+/CD24^-$ *mesenchymal-like* CSCs after antiretroviral drug treatment was supported by the immunofluorescence and the RNA-Seq results. Therefore, the $CD44^+/CD24^-$ *mesenchymal-like* CSCs were isolated by a cell sorter, and their responses to drugs examined. The $CD44^-/CD24^-$ population was used as a non-CSC control since, in theory, they should be easier to culture.

Although the sorted cells could survive and proliferate slowly, the additional stress of drug treatment compromised their viability (in appendix Figure S5-5). Thus, to solve this problem, different staining and sorting conditions were tested, until finally cell viability and operability were increased (expanding the cell numbers to 10^8, reducing the staining to twenty minutes, and increasing the sorting speed). After sorting, the purity of the $CD44^+/CD24^-$ CSC population was approximately 98%, while the purity of the $CD44^-/CD24^-$ population was usually around 40%. This was because the fluorescence signals from the fluorescence positive cells could be masked by other cells resulting in them being incorrectly classified as fluorescence negative cells (Figure 5-7A and 5-7B).

The other problem arising from the separation of CSC types was that some cancer cells frequently switch between CSC and non-CSC status (CSC plasticity). This CSC maintenance challenge has been reported in several papers (Meacham and Morrison, 2013, Liu et al., 2015) and has strongly influenced CSC drug response experiments. After sorting, the $CD44^+/CD24^-$ CSCs differentiated into non-cancer stem cells very rapidly (Figure 5-7D); also, the non-CSC

control ($CD44^-/CD24^-$ cells) was able to convert to the patterns of the total cancer population within only a few days (Figure 5-7E). After culturing sorted cells for five days, the CSC-enriched group (all of the $CD44^+/CD24^-$ cells during sorting) had only about 50% $CD44^+/CD24^-$ CSCs remaining (Figure 5-7D). The proportion of the $CD44^+/CD24^-$ population in the CSC-enriched group was relatively higher than in the $CD44^-/CD24^-$ non-CSC control group (27%) and in non-sorted total cells (29.3%) (Figure 5-7C and 5-7E). Yet the accuracy of the CSC experiments may be highly affected by the presence of the non-CSC cells (approximately 50%) in the sorted CSC-enriched group. Thus, it was necessary to improve the CSC maintenance method before performing further experiments.

The cancer cell line employed in this experiment was MCF10CA1α. A greater proportion of the sorted-CSCs in MCF10AT rapidly differentiated into non-CSCs, one to two days after culturing; whereas the $CD44^+/CD24^-$ cells dominated in the MDA-MB-231 cell line of which more than 85% of cells were the *mesenchymal-like* CSCs. Thus, cell sorting was unnecessary in this cell line preventing further damage to the cells during the staining and sorting process.

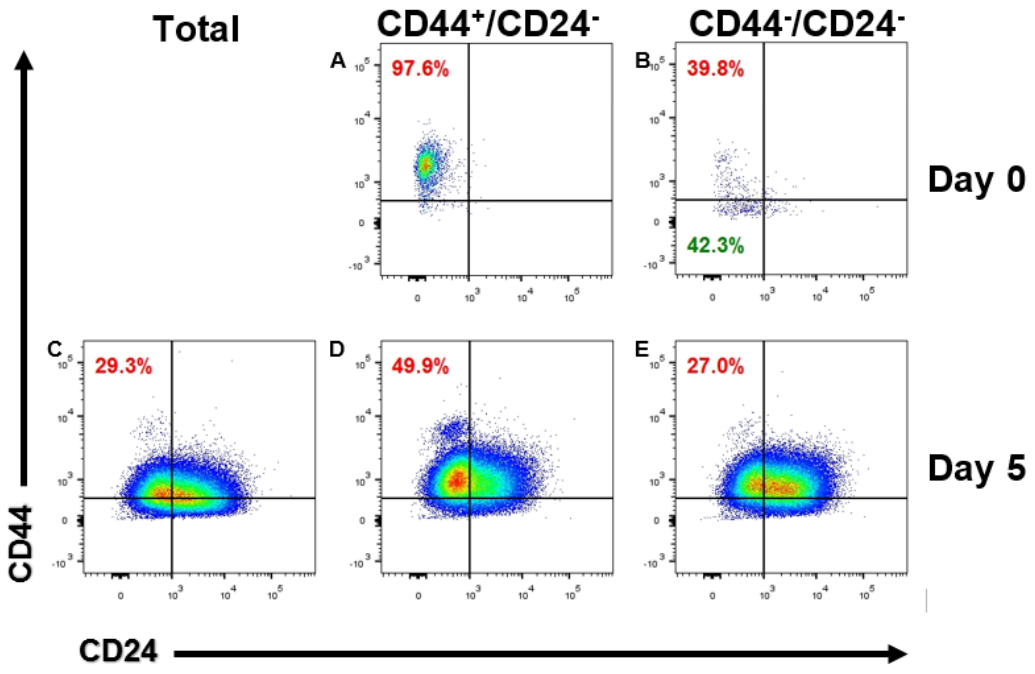


Figure 5–7 The percentage of the $CD44^+/CD24^-$ cells in the unsorted total population and sorted-CSC-enriched population versus the percentage of the $CD44^-/CD24^-$ cells

The dots in the up-left corner of each FACS figure represent the $CD44^+/CD24^-$ *mesenchymal-like* CSCs. The dots in the down-left corner of each FACS figure represent the $CD44^-/CD24^-$ cells which are used as non-CSC control. The frame lines were set based on their unstained controls.

5.2.3.2 Mesenchymal-like CSCs culturing in serum-free medium

Many published studies have suggested that maintaining cancer cells in CSC status is extremely challenging due to a strong tendency of CSCs from epithelial cancers to differentiate into non-CSCs in cultured conditions (Meacham and Morrison, 2013, Liu et al., 2015). However, there are several methods that can potentially support CSC maintenance. They combine cell sorting, chemo- or radiotherapy, mesenchymal induction, and serum-free culturing (Liu et al., 2015). The advantages and disadvantages of these methods are described below.

Firstly, cell sorting can result in a relatively pure CSC population but is unable to provide long-term CSC maintenance. Secondly, some epithelial-mesenchymal transition (EMT) inducers, such as Twist and Snail, may help cells to remain in CSC status; however, adding extra inducers might strongly affect CSC experiments by directly or indirectly altering some CSC markers, thereby potentially compromising the data. Similarly, chemotherapy and radiotherapy may also result in the same problems as the EMT inducers. Lastly, serum-free cultivation alone is unable to greatly enrich CSCs. However, the most feasible method for obtaining stable CSCs could be to combine cell sorting with serum-free cultivation.

Sorted $CD44^+/CD24^-$ CSCs were cultured in serum-free medium in order to investigate the possibility of CSC enrichment. Surprisingly, after five days of culturing, there was no observable CSC population difference between serum-free and serum-containing medium (Figure 5-8B and 5-8C). Culturing sorted-CSCs in the serum-containing medium for eight hours and then changing the medium to serum-free medium for the remainder of the culturing period resulted in the highest percentage (nearly 60%) of $CD44^+/CD24^-$ CSCs (Figure 5-8C). Culturing sorted-CSCs in the serum-containing medium for 48 hours and then changing the medium to serum-free medium maintained a higher percentage (about 53%) of the $CD44^+/CD24^-$ CSCs (Figure 5-8) compared with culturing sorted-CSCs in serum-free medium for the whole culturing period (about 46%). Although none of these culture conditions could maintain stable CSCs for five days, cultivation for 8 hours in serum-containing medium followed by four days in the serum-free medium resulted in the best $CD44^+/CD24^-$ CSC enrichment. Therefore, these culture conditions were used for observing the effects of drugs on $CD44^+/CD24^-$ CSCs.

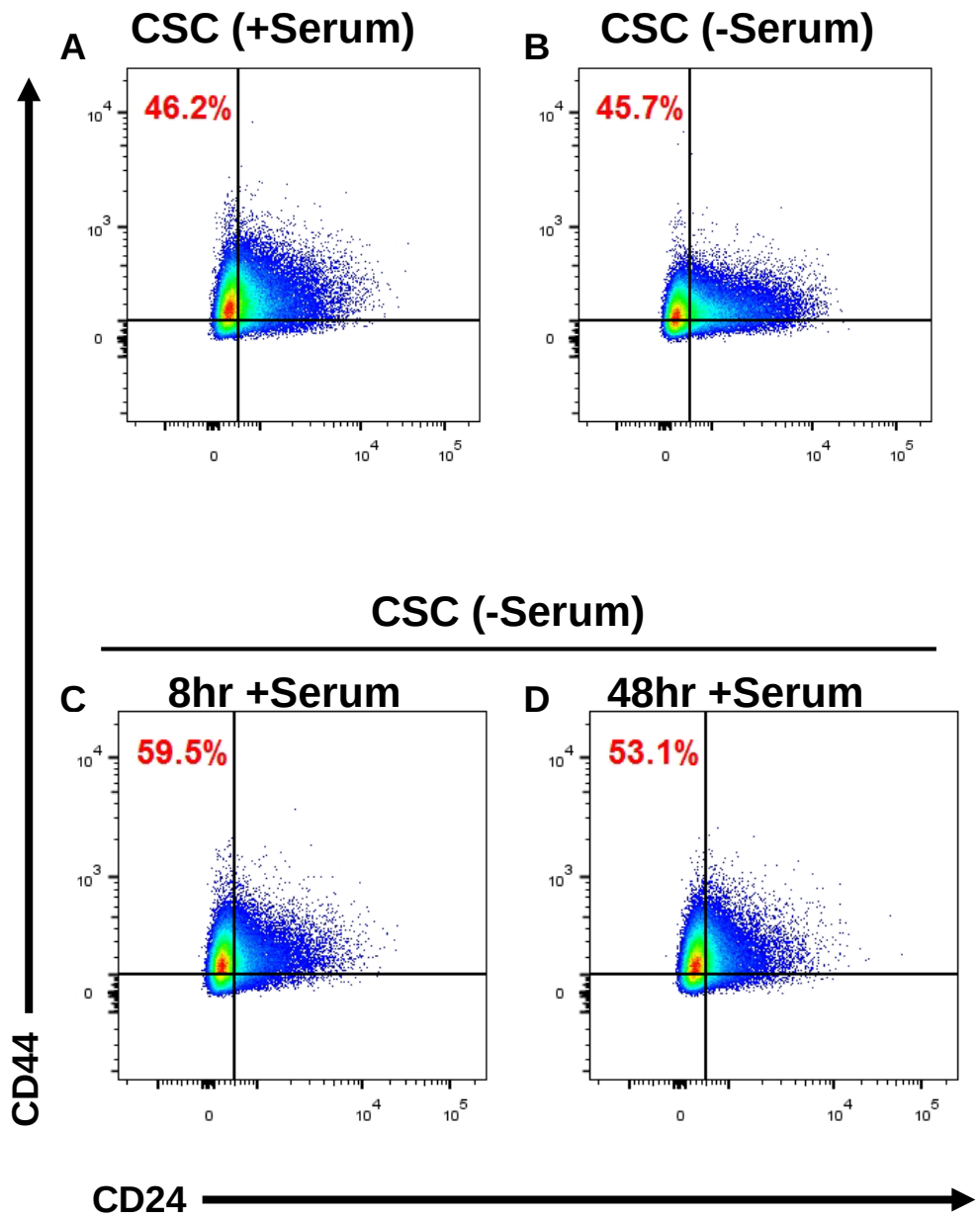


Figure 5–8 Sorted-CD44$^+$/CD24$^-$ CSCs cultured under different serum-free conditions
The dots in the up-left corner of each FACS figure represent the CD44$^+$/CD24$^-$ *mesenchymal-like* CSCs. The frame lines were set based on their unstained controls. (A) Sorted CSCs cultured in the complete medium with 5% horse serum supply for 96 hours. (B) Sorted CSCs cultured in the complete medium without serum supply for 96 hours. (C) Sorted CSCs cultured in the complete medium without serum supply for 8 hours and then exchanged the medium to the complete medium with 5% horse serum supply for 88 hours. (D) Sorted CSCs cultured in the complete medium without serum supply for 48 hours and then exchanged the medium to the complete medium with 5% horse serum supply for 48 hours.

5.2.3.3 Drug treatment of $CD44^+/CD24^-$-sorted CSCs

When sorted-MCF10CA1α-$CD44^+/CD24^-$ cancer cells were cultured in modified serum-free medium, less than half of the $CD44^+/CD24^-$ CSCs differentiated into non-CSCs. About 62.4% of the $CD44^+/CD24^-$ CSCs were observed in the untreated control; while, 45.5% of the $CD44^+/CD24^-$ CSCs were observed in EFV-treated cells (Figure 5-9). Contrary to the results in the total population, the CSC-enriched population showed a reduction in the $CD44^+/CD24^-$ CSC population after EFV treatment. This suggests that EFV could be effective for eliminating the $CD44^+/CD24^-$ *mesenchymal-like* CSCs.

Additionally, the qRT-PCR data (Figure 5-10) suggested that the expression of LINE-1 RNA in MCF10CA1α was reduced both in the CSC-enriched population and in the total population after EFV treatment. However, the low purity of the $CD44^+/CD24^-$ CSCs might affect the results. In theory, with a population of pure CSCs, we should be able to determine whether the drug treatment induces them to transform into non-CSCs. Yet, with a mixture of CSCs and non-CSCs, it was difficult to determine the effect of the drugs on CSCs. The percentage of CSCs in the population not only resulted from CSC transformation but also from non-CSCs transforming into CSCs. Therefore, in reality, the drug effect of CSCs transforming to non-CSCs could be more frequent than the observation showed in the FACS experiments. This suggested that the accuracy of the experiments could be enhanced if the purity of the studied CSCs could be increased. Since maintaining stable $CD44^+/CD24^-$ CSCs was an issue, the tumorsphere CSC functional assay was investigated as a compromise solution.

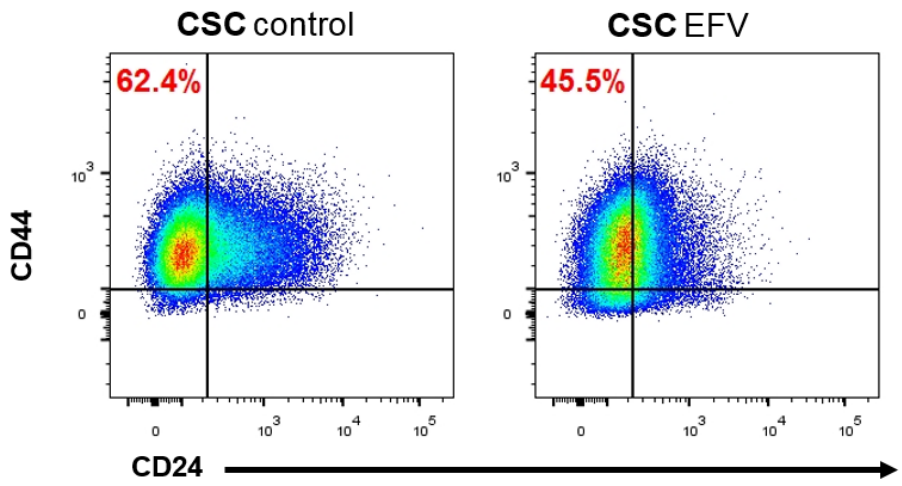


Figure 5–9 Sorted-CD44$^+$/CD24$^-$ CSC after drug treatment
The dots in the up-left corner of each FACS figure represent the CD44$^+$/CD24$^-$ *mesenchymal-like* CSCs. The frame lines were set based on their unstained controls. These sorted *mesenchymal-like* CSCs were cultured in the serum-free complete medium for 8 hours and then changed the medium to the complete medium with 5% horse serum containing DMSO (control) or EFV for 88 hours. The percentage of the CD44$^+$/CD24$^-$ population was lower in EFV treated cells (right-hand side) than in the control cells (left-hand side).

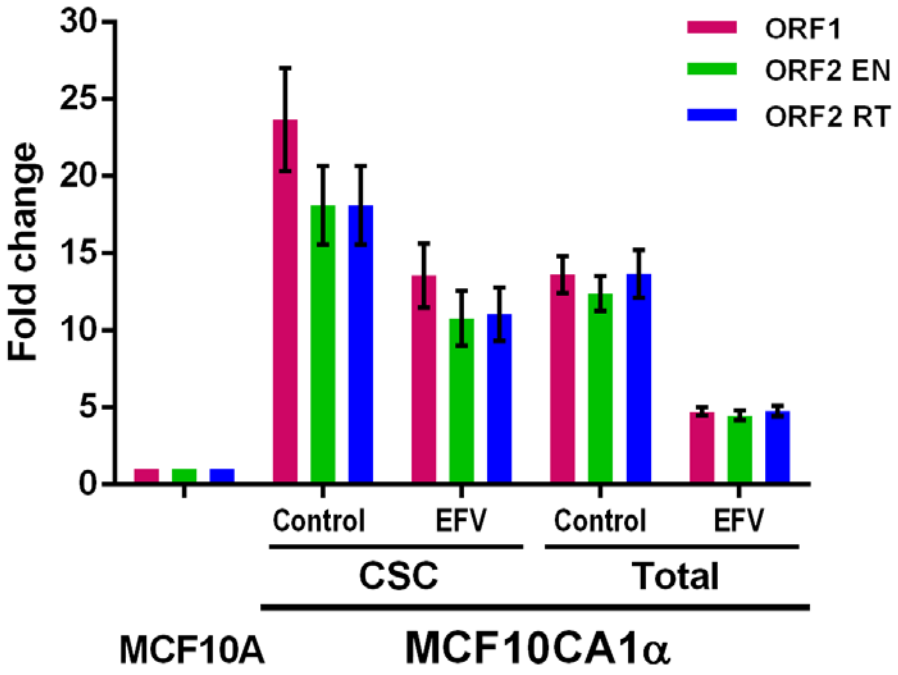


Figure 5–10 qRT-PCR data comparing LINE-1 RNA levels within EFV-treated and untreated MCF10CA1α cells and the CSCs isolated from MCF10CA1α cell lines
The pink bars represent the expression of LINE-1 *ORF1* RNA; the green bars represent the expression of the endonuclease region of LINE-1 *ORF2* RNA; the blue bars represent the expression of the reverse transcriptase region of *ORF2* RNA. Error bars: ±SD, n = 4.

5.2.4 Functional assay: three-dimensional MammoCult culture for tumorsphere formation.

After failing to maintain CSCs under culturing conditions, growing cells in the MammoCult™ Medium (Stemcell technology) was used as an alternative strategy for undertaking CSC experiments. Cultivation of cells in MammoCult™ Medium is a three-dimensional culture method specifically used for breast cancers and only cells with CSC ability can survive and proliferate under these conditions. It has been used for culturing many breast cancer cell lines, including MDA-MB-231, MCF7, SKBR3, AU565, BT474, and SUM149 which then progress to form tumorspheres. These tumorspheres can subsequently form tumours if injected into immunodeficient mice (Kim and Alexander, 2014), indicating that the cells which can grow in MammoCult™ Medium are functional CSCs. However, there are no reports showing that the MammoCult™ Medium has been used for the MCF10 cell series.

A preliminary study was performed to observe whether MCF10A, MCF10AT, and MCF10CA1α cells could be cultured in MammoCult Medium. After transferring MCF10A, MCF10AT, and MCF10CA1α cells from two-dimensional to three-dimensional cultures with MammoCult Medium, cells started to form spheres which grew over subsequent days. On the third day, cells started to form complete tumorspheres; and on the fourth day, MCF10AT and MCF10CA1α tumorspheres formed dark centres which were difficult to split (Figure 5-11). MCF10A formed small and loose tumorspheres within two days; whereas, MCF10AT and MCF10CA1α formed small and tight tumorspheres within three days. Therefore, the third day of the MammoCult culture (the solid tumorsphere formed) was chosen as the starting point for the drug treatment.

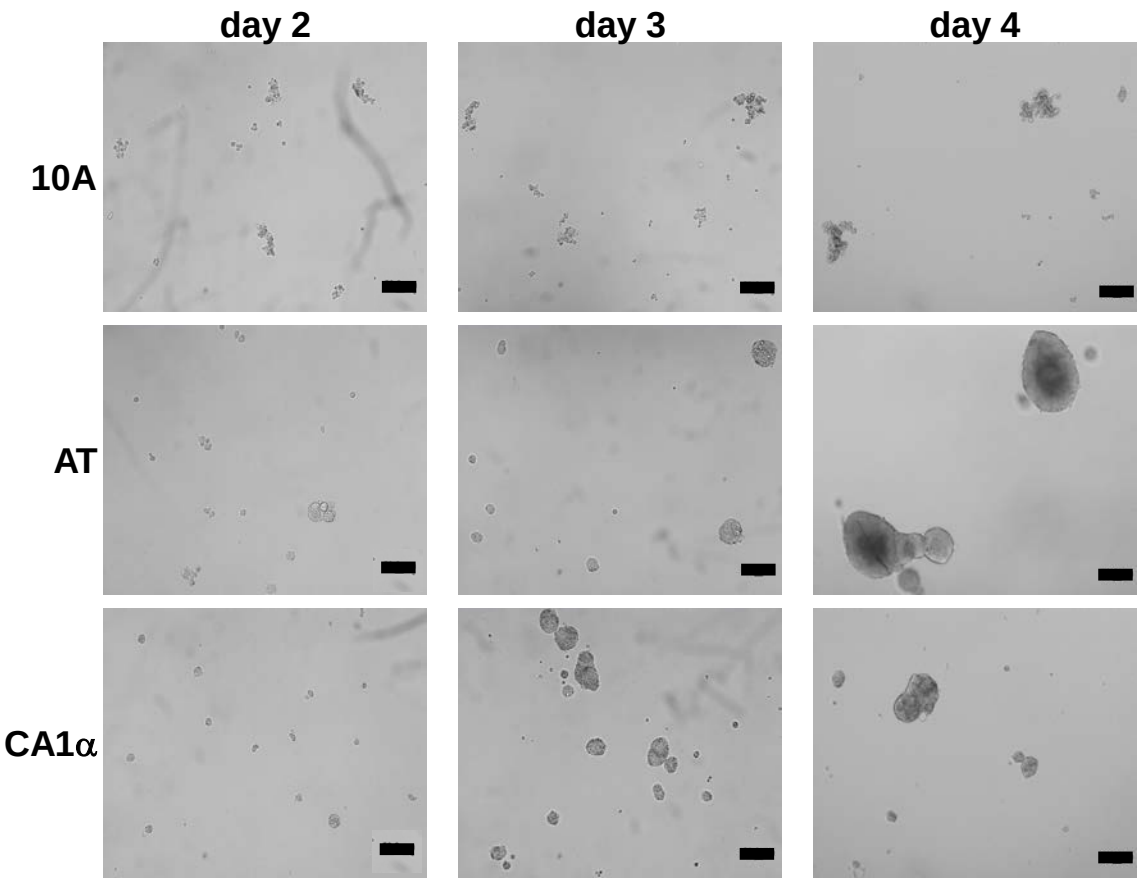


Figure 5–11 Pre-test for culturing cells using the MammoCult three-dimensional culture conditions

Images were taken every day from day 2 to day 4 after seeding in order to determine a suitable MammoCult culture condition for MCF10A, MCF10AT, and MCF10CA1α cell lines. Scale bar: 200 µm.

5.2.5 The influence of Efavirenz on tumorsphere formation

Interestingly, the relative numbers of tumorspheres formed for each cell line mostly correlated with the relative malignancy of the cancer cell line. The most malignant cell line, MDA-MB-231, formed the largest number of tumorspheres compared with other cell lines; while the low-invasive MCF10AT produced the lowest number of tumorspheres. BT-549 was an exception: this highly invasive cell line produced relatively few tumorspheres compared with the low-invasive BT-20 cell line; however, this cell line did form larger tumorspheres (larger than 400 μm) than all the other cell lines.

Treating MCF10A, MCF10AT, MCF 10 CA1α, BT-549, BT-20, and T47D cells with EFV or SPV for seven days, resulting in a reduction in the size of the tumorspheres. In addition, the structures of the tumorspheres were altered from tight to loosen (Figure 5-12). It was also noted that the number of tumorspheres decreased after EFV treatment (Table 5-4, Figure 5-13).

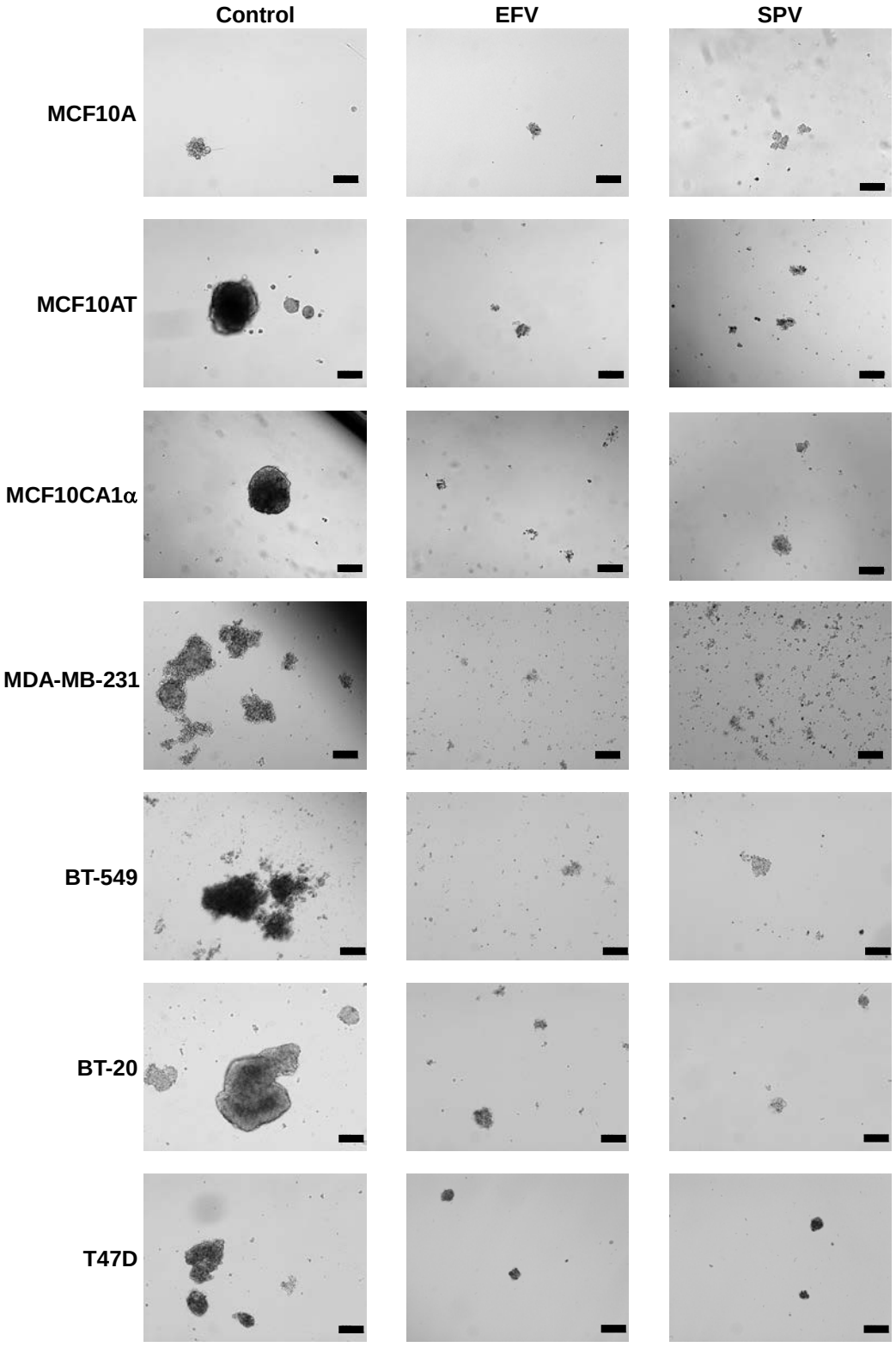


Figure 5–12 MammoCult culture with drug treatments for seven days
Harvested MCF10A, MCF10AT, MCF10CA1α, MDA-MB-231, BT-549, BT-20, and T47D cells from 2D cultures, then seeded an appropriate number of live cells in the complete MammoCult Medium. After

three days of sphere formation, the cells were incubated with EFV and SPV for seven days. Scale bar: 200 μm.

After EFV treatment, the total number of tumorspheres decreased from 92.3±16.2 to 11.0±6.2 in MCF10AT; from 198.0±56.8 to 18.3±18.0 in MCF10CA1α; from 422.0±17.8 to 8.3±8.5 in MDA-MB-231; from 122.0±24.7 to 47.3±38.0 in BT-549; from 211.0±24.0 to 49.0±30.4 in BT-20; and from 158.0±36.9 to 10.8±4.5 in T47D (Table 5-4). There were significant differences in the total number of tumorspheres for untreated- and EFV-treated cells in all cancer cell lines tested (in appendix Table S5-1). Notably, after EFV treatment, the tumorsphere number was 50 times less than the control in MDA-MB-231; 14.5 times less in T47D; 10 times less in MCF10CA1α; and 8 times less in MCF10AT. Although the tumorsphere number was only reduced 1.5 times in BT-549 after EFV treatment, it showed the greatest change with 32 tumorspheres larger than 400 μm being reduced to no tumorsphere larger than 400 μm (Table 5-4).

The data suggested that tumorspheres can be eliminated by EFV and SPV treatment, thus indicating that EFV and SPV might be able to treat both non-CSCs and CSCs.

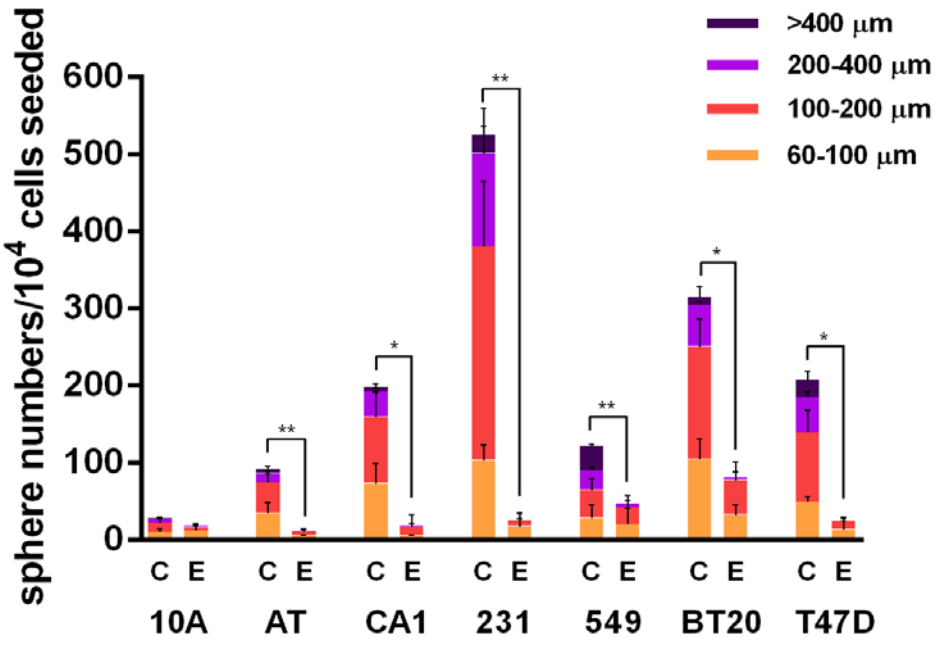


Figure 5–13 Numbers of spheres being formed after EFV treatment in MCF10A, MCF10AT, MCF10CA1α, MDA-MB-231, BT-549, BT-20, and T47D cell lines

This figure is a quantitative bar chart for the tumorsphere formation results. The dark-purple bars represent the spheres larger than 400 μm; the light-purple bars represent the spheres larger than 200 μm but smaller than 400 μm; the dark-orang bars represent the size of the sphere in between 100 μm to 200 μm; the light-orange bars represent the size of the sphere in between 60 μm to 100 μm. Any clump smaller than 60 μm is not considered as tumorsphere. Two-tailed, paired Student's *t*-test was conducted for statistical analysis, the total cell numbers are compared between untreated- and EFV-treated conditions. *p<0.05; **p<0.01. Error bars: ±SD, n = 3. C: untreated-control cells, E: EFV-treated cells.

Table 5-4 The numbers of tumorspheres formed after seven days of EFV treatment.

Number of spheres /10^4 cell seeded		60-100 μm	100-200 μm	200-400 μm	>400 μm	total
MCF10A	control	9.5±4.5	12.3±4.7	5.0±2.7	1.0±0.0	27.8±11.7
	EFV	11.5±9.2	4.5±2.2	1.8±1.9	0.3±0.6	18.2±10.3
MCF10AT	control	34.0±14.1	39.7±11.0	12.3±2.3	6.3±3.1	92.3±16.2
	EFV	5.3±1.2	4.3±4.0	1.3±2.3	0.0±0.0	11.0±6.2
MCF10CA1α	control	73±25.5	86.3±31.3	32.3±5.0	6.7±3.2	198±56.8
	EFV	5.0±1.0	11.7±15.0	1.7±2.9	0.0±0.0	18.3±18.0
MDA-MB-231	control	103.0±20.0	277.0±85.2	121.0±34.9	24.0±34.1	422.0±17.8
	EFV	17.3±17.2	7.7±8.6	0.7±1.2	0.0±0.0	8.3±8.5
BT-549	control	28.0±16.5	36.0±15.1	25.7±4.7	32.0±2.0	122.0±24.7
	EFV	19.7±21.8	22.0±15.1	5.7±3.2	0.0±0.0	47.3±38.0
BT-20	control	104.0±26.7	146.0±36.1	54.7±4.0	10.7±12.4	211.0±24.0
	EFV	32.0±13.1	45.0±24.1	4.0±6.9	0.0±0.0	49.0±30.4
T47D	control	48.8±7.2	90.0±29.0	45.7±7.4	22.3±10.8	158.0±36.9
	EFV	13.3±14.6	10.7±4.7	0.2±0.3	0.0±0.0	10.8±4.5

5.3 Discussion

CSCs are believed to be one of the major causes of epithelial cancer metastasis, patient relapse, and therapy resistance. Therefore, specifically targeting CSCs may potentially enhance prognostic outcomes for patients. However, the lack of universal CSC markers increases the difficulties of CSC research. In this study, a few commonly used methods for identifying BCSC were utilised to further understand the effects of the antiretroviral drugs on BCSCs.

BCSCs which potentially produce drug resistance in breast cancers can be eliminated by antiretroviral drugs. According to the evidence gained from this study, the percentage of *epithelial-like* CSCs decreased after the drug treatment, suggesting that antiretroviral drugs can effectively reduce BCSCs. Although *mesenchymal-like* CSCs experiments revealed opposite results, mesenchymal CSC-enrichment experiments can partially explain this discrepancy. The results obtained from the tumorsphere functional CSC assay further supported our finding of the effectiveness of antiretroviral drugs at eliminating CSCs. Therefore, the drugs may potentially be a powerful tool for treating TNBCs including CSCs and non-CSCs. However, there are some limitations of the CSC study which are yet to be overcome.

The major limitation of hindering CSC research is how best to identify CSCs. To date, there is no single method which has been proven to identify all CSCs. Some studies only focus on one type of CSCs (Sheridan et al., 2006, Fillmore and Kuperwasser, 2007, Jaggupilli and Elkord, 2012); while other studies have investigated different types of CSCs and observed similar drug responses (Mishra et al., 2017, Lin et al., 2013). Yet, according to the antiretroviral drug treatment experiments described in our study, different types of CSCs respond differently to different drug treatments. Antiretroviral drugs reduced the proportion of the $ALDH^{high}$ *epithelial-like* CSCs, suggesting that antiretroviral drugs can effectively eliminate the *epithelial-like* CSCs in breast cancers. In contrast, the percentage of the $CD44^{+}/CD24^{-}$ *mesenchymal-like* CSCs increased after drug treatment in all TNBC cell lines tested. The LINE-1 shRNA direct inhibition experiment supported this result implying that LINE-1 inhibition may potentially enlarge the *mesenchymal-like* CSC population within TNBCs. This antiretroviral drug-induced CSC enrichment pattern could be caused by the drug-resistant ability of CSCs; however, the exact regulation mechanisms remain unclear. These conflicting

results show the heterogeneity amongst CSCs and highlight the challenges facing CSC researchers.

Another significant limitation to CSC research is CSC maintenance. It is difficult to preserve sorted CSCs with their CSC status intact. In traditional two-dimensional cell culture, some CSCs always tend to differentiate into non-CSCs in order to retain the original composition of cancers (Tang, 2012, Liu et al., 2015). This phenomenon has also been observed in this study, hence our difficulties with the BCSC enrichment experiments. Several potential solutions were considered, but only cell sorting in combination with serum-free medium seemed to enrich the *mesenchymal-like* CSC population. Even though this method managed to only maintain about 60% of breast cancer cells in CSC status, a decreased proportion of the *mesenchymal-like* CSCs was observed after EFV treatment. This experiment indicated the ability of antiretroviral drugs to reduce BCSCs; however, more solid evidence is needed to confirm this result. Therefore, currently, a widely used BCSC enrichment assay, the tumorsphere BCSC functional assay was employed to create BCSC enrichment conditions which allowed us to examine the effectiveness of antiretroviral drugs at reducing BCSCs. The issue of CSC maintenance underlined CSC's plasticity and this was a driving motivation for investigating which anticancer drugs can eliminate both CSCs and non-CSCs.

CSC treatment has been drawing great attention in cancer research. CSCs have stronger drug tolerance than non-CSCs do, thus they are very difficult to deal with. Additionally, some cancer cells might be able to switch between CSC status and non-CSC status; however, the mechanisms involved are still not understood. A drug which can kill both CSCs and non-CSCs is expected to be superior to any other currently used in clinical practice. This study addressed the possibility of repurposing antiretroviral drugs as a novel anticancer therapy and highlighted their potential for inhibiting the growth of CSCs.

Chapter 6 Conclusions and discussions

6.1 Summary

The main goal of this project was to determine whether antiretroviral drugs can be repurposed for treating triple-negative breast cancers (TNBCs). This cancer cell subtype is difficult to treat with the limited therapeutic strategies currently available. The initial part of this study determined the effectiveness of commonly used antiretroviral drugs, as anticancer agents, in TNBCs. The second part of this study investigated the mechanisms involved in the antiretroviral drug-induced cell responses. The final part of this study examined the effectiveness of antiretroviral drugs in reducing BCSCs which are known to be responsible for many unfavourable prognostic outcomes, including cancer metastasis, drug resistance, and cancer recurrence. A discussion of the outcomes and implications follows.

A significant result arising from this study has been the anticancer effects of two antiretroviral drugs Efavirenz (EFV) and SPV122 (SPV) on TNBCs via LINE-1 inhibition. Out of the six antiretroviral drugs tested, only EFV and SPV strongly decreased cell viability in the TNBC cell lines MCF10AT, MCF10CA1α, MDA-MB-231, BT-549, and BT-20; whereas, no obvious changes were observed in non-cancer control MCF10A cells. The studies have shown that antiretroviral drugs induce cancer cell toxicity and retard cell proliferation in TNBCs. This indicates that the antiretroviral drugs EFV and SPV act specifically and effectively to eliminate TNBCs. We also characterised the morphological and physiological effects of antiretroviral drugs on TNBC cell lines and found that they reduced cell proliferation, suppressed DNA synthesis within the cell cycle, increased programmed (apoptosis) and unprogrammed (necrosis) cell death, and altered cancer cell morphology. Moreover, inhibition of LINE-1 by antiretroviral drug in TNBC cell lines was confirmed by western blot and LINE-1 simulation experiments. Since LINE-1 overexpression has been strongly linked with breast cancer (Patnala et al., 2014), and the function of LINE-1 reverse transcriptase (RT) is reduced by antiretroviral drugs (Dai et al., 2011), we have established a link between antiretroviral drugs reducing cancer cells and inhibition of LINE-1. This suggests that using antiretroviral drugs to block LINE-1 RT may be of potential therapeutic value in TNBCs.

The other important finding of this study was that antiretroviral drugs exert their anticancer effects by down regulating fatty acid metabolism. In this component of the project, we used whole-genome RNA sequencing (RNA-Seq) analysis to further our understanding of the possible anticancer mechanisms involved in MCF10AT, MCF10CA1α and MDA-MB-231 cell lines. According to the RNA-Seq data analysed by the STRING-DB protein-protein interaction search tool, several key genes involved in fatty acid metabolism play a key role in antiretroviral drug treatment. These genes included stearoyl-CoA desaturase (*SCD*), fatty acid synthase (*FASN*), and acyl-CoA synthetase long-chain family (*ACSL*) which were all significantly down regulated after the drug treatments. Since fatty acid metabolism has been strongly connected with cancer development, its down regulation makes it highly relevant to the antiretroviral drug-induced anticancer activity. Thus, investigation of the relationship between LINE-1 inhibition and suppression of fatty acid metabolism could be an area for future study.

The fatty acid metabolism pathway was not the only pathway implicated in antiretroviral drug treatment in the STRING-DB analysis. Other pathways such as 'FoxO signalling pathway', 'microRNAs in cancer' pathway, and 'Amino sugar and nucleotide sugar metabolism' pathway were also highlighted. These pathways may help us to further understand the mechanisms involved in the molecular and morphological changes which were observed during the drug treatment. Interestingly, some of the pathway have been correlated with anticancer effects, but others have been associated with unfavourable prognostic outcomes of cancers. These diverse outcomes further emphasise the complexity of TNBCs as illustrated by the varied phenotypes encountered for a range of antiretroviral drug-treated TNBC cell lines.

The increasing expression of 'microRNAs in cancer'-related genes in the drug-treated TNBCs appeared to be associated with enhancing cancer cell malignancy. Genes involved in the 'microRNAs in cancer' pathway, such as CD44, have been strongly linked with poor prognostic outcomes in breast cancers (Hu et al., 2018). In addition, not only the genes associated with cancer-related microRNAs, but also the cancer-related microRNAs, such as *miR-21*, were upregulated after drug treatment. We observed significant increases of *miR-21* in antiretroviral drug-treated TNBC cell lines. An increase in the expression of *CD44* and *miR-21* in cancer cells is more than a sign of cancerous predisposition, it can also be associated with cancer stem cell

(CSC) enrichment. Therefore, this led us to consider how CSCs could be involved in antiretroviral drug treatment.

The ANOVA analysis of RNA-Seq data also strongly suggested that CSCs may be induced during antiretroviral drug treatment. Quite a few CSC-related genes were highlighted in the RNA-Seq analysis with some decreasing and others increasing upon drug treatment. For instance, two fatty acid metabolism related-genes *SCD* and *FASN*, which had been associated with CSC (Kuo and Ann, 2018), were down regulated as a result of drug treatment. In contrast, *CD44* and *PROCR*, two well-known CSC indicators (Hwang-Verslues et al., 2009), were up regulated. These inconsistent results can at least partially explain the dual directions of cancer morphological change in the drug-treated MCF10AT and MCF10CA1α cells. They also highlight the importance of researching antiretroviral drug responses in CSCs.

Antiretroviral drugs target both general cancer cells and CSCs in TNBCs. In our experiments, different types of BCSCs responded differently to antiretroviral drugs. The high aldehyde dehydrogenase (ALDH) activity epithelial-like CSCs were significantly reduced after antiretroviral drug treatments; whereas, the $CD44^{+}/CD24^{-}$ mesenchymal-like CSCs showed an increase. These results may explain the diametrically opposed results observed in previous experiments: some of the antiretroviral drug-treated TNBCs presented mesenchymal and CSC markers associated with unfavourable prognostic outcomes. These findings highlight the complexity and the heterogeneity of TNBCs and further indicate that different types of CSCs may have varying degrees of drug-resistance. The mesenchymal-like CSCs probably have greater resistance to antiretroviral drugs than other cancer cells. Finally, the tumorsphere CSC functional assay demonstrated that CSCs can be eliminated by antiretroviral drugs, thus continuous antiretroviral drug treatment at an appropriate concentration may be able to reduce both non-CSCs and CSCs simultaneously.

In summary, several significant outcomes have been achieved during the course of this project. The most significant finding to emerge was that EFV and SPV could potentially be valid anticancer drugs for treating TNBCs. They are effective against CSCs and non-CSCs by their regulation of the fatty acid metabolism pathway. Although future studies are required to directly link LINE-1 inhibition and cancer fatty acid metabolism, this study suggests that LINE-1 is a potential therapeutic target in the treatment of TNBCs. However, considering the ability of

CSCs to resist existing drug treatment regimens (Singh and Settleman, 2010), a thorough evaluation is necessary to assess the efficacy of antiretroviral drugs as an anti-cancer therapy. More experiments are warranted so that we can fully understand how antiretroviral drugs act as inhibitors of cancer.

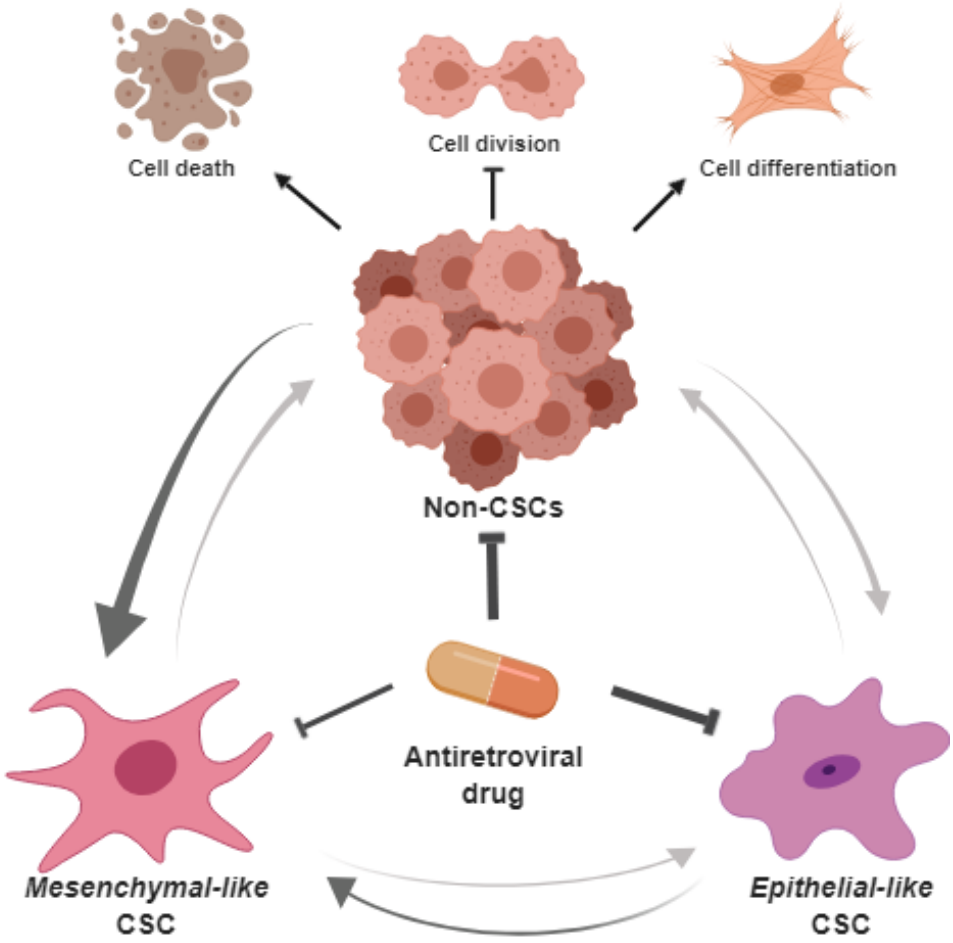


Figure 6-1 putative functions of antiretroviral drugs in TNBC treatment
Antiretroviral drugs may antagonise the progress of TNBCs by targeting both CSCs and non-CSCs. They can promote cell differentiation in undifferentiated cancer cells, can suppress cancer cell division, and can induce cancer cell death. *Mesenchymal-like* CSCs are more resistant to antiretroviral drugs compared with other types of TNBC cells, although the drugs can also reduce the *mesenchymal*-like CSCs. In this cancer reducing process, cancer cells may switch their status between CSCs and non-CSCs because of CSC plasticity. TNBC cells tend to have a preference for maintaining the *mesenchymal-like* CSC status under antiretroviral drug treatment condition. However, the functional CSC assay proves that CSCs can also be targeted by the drugs. Thus, antiretroviral drugs should be considered as a potential therapeutic strategy for treating TNBC patients.

6.2 Limitations

Despite LINE-1's long-time association with epithelial cancers, the nature of this association has remained ambiguous. LINE-1 is strongly inhibited in normal cells through many different endogenous mechanisms. Thus, the two major tools currently used in protein functional studies, a gain of function (overexpression) and a loss of function (inhibition), might be less applicable in LINE-1 research. The overexpression of LINE-1 in normal cells expressing low levels of LINE-1 would require the cell's self-defence mechanisms to be overcome, which would be a challenge. Although a modified LINE-1 sequence can partially reduce the influence of such self-defence mechanisms, it may have a different structure or function to the original LINE-1. We must also remember that the exact protein structure of LINE-1 ORF2 protein, the predicted target of antiretroviral drugs, is still unproven. Whether antiretroviral drugs can directly target LINE-1 RT also remains unclear. Until now, no direct link has been confirmed between antiretroviral drug treatment, LINE-1 inhibition, and cancer cell elimination in TNBCs. However, our results suggest that LINE-1 expression is correlated with cancer growth in TNBCs. Thus, there is a great opportunity for LINE-1 to be advanced as a potential therapeutic target in TNBCs.

Another limitation of this study concerns the efficacies of antiretroviral drugs. It is still unclear whether the diverse drug-induced morphological changes are caused by insufficient drug treatment, possibly due to the characteristics of the drugs and drug resistance of cancers. EFV is a commercially available antiretroviral drug with a long history of treating HIV patients. No reports have been published regarding its cell permeability and stability (Bastos et al., 2016). Such an issue may also apply to the newly synthesised SPV which may be unstable. We found that the efficacy of SPV decreased continuously with time and a precipitate was formed after two years storage in a -80°C freezer. The stability of SPV would have to be improved before any further experiments are conducted in the future. As for drug resistance, EFV has been recognised as an inhibitor of the breast cancer resistance protein (BCRP) which is responsible for a great proportion of multidrug resistance in breast cancers (Weiss et al., 2007, Peroni et al., 2011). This implies that antiretroviral drugs may reduce cancer drug resistance and operate within the target cancer cells. It is likely that the diverse molecular and morphological changes in the drug-treated TNBCs might be caused by different cell responses to wild-type p53

suppression and mutant p53 suppression, as well as the involvement of CSCs, rather than by efficacy of the drug. However, more evidence is needed to test these assumptions.

CSC research is an evolving field of cancer biology which is facing several challenges, including CSC heterogeneity and CSC plasticity. These characteristics of CSCs mean that it is difficult to assess the outcomes of drug treatment. A primary area of concern in the CSC field is the lack of universal CSC markers. Several molecular markers, such as $CD44^+/CD24^-$, and functional markers, such as high ALDH activity, have been recognised as CSC markers. However, only a small amount of different types of CSCs overlap. Thus, which types of CSCs can represent the major CSC population in cancers is currently under debate. In addition, different types of CSCs respond differently to different drugs, making it increasingly difficult to determine effective treatments. Some cancer cells even have the ability to shift between CSC and non-CSC status, the triggers and mechanisms of which are unclear. These characteristics greatly affect CSC maintenance and reduce the accuracy of CSC-enrichment experiments. Moreover, traditional in vitro two-dimensional culture conditions might not reproduce tumour microenvironments. This raises questions of whether the experimental conditions actually mimic CSCs' native environment. Although the MammoCult three-dimensional culture may partially address this problem, little is known about how many true CSCs can be accumulated in this culture. CSC research is a relatively young and underdeveloped field of cancer research, many hypotheses and assumptions need to be tested. Further investigation of this discipline will enable more effective therapeutic methods to be developed for cancer patients.

The first evidence of cancer stem cells (CSCs) arose in human acute myeloid Leukaemia about 25 years ago (Lapidot et al., 1994); however, so far, there have been no specific markers for detecting CSCs. According to the current literature, $CD44^+/CD24^-$ and $ALDH^+$ are widely used and recognised as CSC markers in breast cancer (Liu et al., 2014). Nevertheless, the risk of exclusively relying on these CSC markers should also be considered. For instance, miR-21 is believed as a valid CSC marker which inhibits tumour suppressor genes such as PTEN and further induces cancer proliferation and migration (Han et al., 2011; Buscagla and Li, 2011). Interestingly, in our RNA-Seq data, *pten* increases after antiretroviral drug treatment, although miRNA level is relatively high in drug-treated TNBCs. This unexpected result has also been found in Frankel's group (2008). They have confirmed that miR-21 inhibitors were very

unlikely to alter PTEN expression in breast cancer suggesting cell- and tissue-specific effects may influence the exact functions of these “markers” (Frankel et al., 2008). Therefore, it is dangerous to ignore the risk of using specific markers for determining cell status, supportive evidence is always essential for confirming the results.

6.3 Future research

This study revealed the anticancer and anti-CSC effects, as well as the possible mechanisms, of antiretroviral drugs in TNBCs; however, additional experiments are needed to validate and expand these findings.

Based on the RNA-Seq results, supplementary studies are feasible. RNA-Seq analysis suggests that several signalling pathways and genes were involved in antiretroviral drug treatment. Some of them have been linked to anticancer effects; others to the increase of unfavourable prognostic outcomes. A clear link between antiretroviral drug treatment and mitochondria has been shown in the RNA-Seq network analysis. Many genes associated with mitochondria, especially with mitochondrial complex I, have been highlighted in the network analysis of TNBC+MCF7 gene set. Mitochondria may play an important role in tumorigenesis, but can be damaged by antiretroviral drugs through oxidative stress (Hecht et al., 2018). A well-known repurposed anticancer drug, Metformin, has been shown to reduce tumour growth by inhibiting mitochondrial complex I and is associated with oxidative stress (Wheaton et al., 2014). Metformin has also been connected with mitochondrial enlargement which could be responsible for cancer cell death in MCF7 (luminal A) breast cancer (Zhuang and Miskimins, 2011). Metformin treatment shows quite a few common phenotypes with antiretroviral drug treatment. Consequently, antiretroviral drugs may induce their anticancer effects in TNBCs by disrupting mitochondrial regulation.

As interesting as some of the results in this study are, it is worth making some improvements to some of the research methods and materials used in this study, in order to address some of the shortcomings of the LINE-1 experiments in TNBCs. Several issues relating to LINE-1 research have not been answered as yet; i.e., the changes in LINE-1 overexpression and inhibition. Using novel techniques such as the CRISPR/Cas9 tool may increase the effectiveness and efficiency of plasmid insertion, and hence increase the success rates of the experiments. The discovery of appropriate TNBC model cells which are easily manipulated would improve the experiments likewise. Moreover, exploring the LINE-1 protein structure would build up a cornerstone and help bridge the gaps in LINE-1 knowledge. Once the exact role of LINE-1 in TNBCs is known, new therapeutic methods can be developed.

Considering the limitations of CSC research discussed previously, several studies could be implemented to expand the comprehension of CSCs in TNBCs. The tumorspheres used in this study were used sparingly because of the restrictions of materials. This was particularly problematic since only a few CSCs remained after antiretroviral drug treatment. Hence it was very difficult to make a comprehensive analysis of any molecular change in the drug-treated cells. Large scale tumorsphere formation is essential for collecting enough cells to allow further investigation of the molecular changes in TNBC's functional BCSCs. These could enhance our understanding of how LINE-1 might affect BCSCs and how the proteins, RNAs, and regulation of microRNAs might be involved. Undoubtedly, more effort is needed to extend our knowledge of CSCs for fighting cancer.

Potential networks between different EFV-induced anti-cancer pathways

Although antiretroviral drugs, especially Efavirenz, have been considered as prospective anti-cancer treatments in various cancers (see introduction section), the exact mechanism that underpins their mode of action is still unclear. In Chapter 4, a potential link to Efavirenz treatment (fatty acid metabolism inhibition and TNBC reduction) was highlighted; however, the detail surrounding their involvement is uncertain. A number of Efavirenz treatment studies performed in different tissues and cell lines indicated alternative pathways including the cannabinoid system (Hecht et al., 2013), oxidative stress(Hecht et al., 2018), Type-I interferon response (De Cecco et al., 2019), and fatty acid metabolism (in this thesis project). Investigating the possible relationships between these pathways may perhaps help us to understand a more comprehensive picture of Efavirenz-induced anti-cancer effects.

Abnormal mitochondria may play a direct and/or indirect role in fatty acid metabolism reduction in antiretroviral drug-treated TNBCs. Mitochondria is the energy factory in eukaryotic cell; loss of its functions may result in lipids, proteins, and DNA damage (Schieber and Chandel, 2014) and this damage may lead to cell death and cell retardation proliferation (Sharma et al., 2011). Some aberrant changes in mitochondria including mitochondrial membrane depolarization (Hecht et al., 2018) and mitochondrial gene alteration (described in Chapter 4) have been observed in antiretroviral drug-treated cancer cells. More specifically, irregular mitochondria can lead to a dysfunctional tricarboxylic acid cycle (TCA cycle) and therefore influence the production of Acetyl-CoA which is crucial for fatty acid metabolism

(Menendez and Lupu, 2017b, Spinelli and Haigis, 2018). Mitochondrial dysfunction may also affect reactive oxygen species (ROS) expression and further alter PI3K/AKT regulations, subsequently reducing fatty acid synthase (FASN) expression in cancers (via Sterol regulatory element-binding protein 1) (Flavin et al., 2010). It seems very likely that abnormal mitochondria could be one of the reasons for Efavirenz-induced fatty acid metabolism inhibition. But how can Efavirenz promotes mitochondrial dysfunction remains a question.

Pathways underpinning Efavirenz-related mitochondrial regulation are still lack of unclear, although evidence suggests that cannabinoid receptors may play a role. Cannabinoids have been reported as anti-cancer agents in different types of cancers by regulating cancer proliferation, invasion, and metastasis (Guindon and Hohmann, 2011, Chakravarti et al., 2014). While the anti-cancer effects of Efavirenz appear to be relevant to cannabinoid receptor expression in certain cell lines (Hecht et al., 2013) and the anti-cancer responses of cannabinoids relate in part to abnormal mitochondria-induced anti-cancer responses making it worthwhile exercise to inspect whether there is any connection in between. More specifically, highly expressed cannabinoid receptors such as cannabinoid receptor 1, cannabinoid receptor 2, and G-protein coupled receptor 55 are considered as a marker of Efavirenz-sensitive cancer cells (Hecht et al., 2013); whereas, overexpressed cannabinoid receptors have been detected in certain types of breast cancers including TNBCs (Chakravarti et al., 2014). This suggests that the anti-TNBCs effects of Efavirenz could potentially be associated with cannabinoid receptor regulation. A review article has highlighted some functional links between cannabinoid receptors and many regulatory pathways in TNBCs (Chakravarti et al., 2014). Among these pathways, ROS pathway and PI3K/AKT pathway, which can be altered by mitochondrial dysfunction are distinguished as downstream events of targeting cannabinoid receptors (Chakravarti et al., 2014). Thus, it is possible that Efavirenz affects cannabinoid receptors resulting in abnormal mitochondria, and further inhibit fatty acid metabolism by blocking Acetyl-CoA synthesis and altering the PI3K/AKT pathway ultimately reducing TNBCs. Interestingly, no direct binding has been detected between Efavirenz and any cannabinoid receptor (Gatch et al., 2013); only the synergistic effect has been reported in Efavirenz and cannabinoid agonists treated cancer cells (Hecht et al., 2013). Hence, more evidence is essential for confirming the role of cannabinoid receptors in Efavirenz-induced anti-cancer effects.

Other than the potential cannabinoid receptor-mitochondria-fatty acid metabolism pathway described above, type-I interferon (IFN-1) can also be a regulator of Efavirenz-induced anti-cancer effects. A possible network among LINE-1, IFN, fatty acid metabolism, and cancer stem cells can be inferred from several indirect lines of evidence. Firstly, inhibiting LINE-1 by antiretroviral drugs can reduce IFN-1 expression in ageing cells (De Cecco et al., 2019) and IFN-1 controls fatty acid metabolism via STAT-regulated mitochondrial TCA cycle in immune cells (Wu et al., 2016a, Raniga and Liang, 2018). Moreover, the expression of INF-1, particularly INF-β, and STAT have been associated with cancer stem cell regulation in TNBCs (Doherty et al., 2017). Therefore, it seems logical to link this information together to propose a pathway involving LINE-1, IFN-1, STAT, mitochondrial TCA cycle, fatty acid metabolism, and TNBC reduction. However, relevant experiments performed in TNBCs are rare, thus further experiments need to be conducted in TNBCs to further prove this hypothesis.

In summary, two possible Efavirenz-induced anti-cancer pathways are proposed in this section, one includes the cannabinoid pathway, the other includes the INF/STAT signalling pathway. Altering these two pathways might be able to drive mitochondrial dysfunction thereby decreasing fatty acid metabolism which could be a key factor in the reduction of TNBCs. Detail experiments are required to confirm exact pathways involved in Efavirenz-induced anti-cancer effects in order to add to our knowledge of antiretroviral drug treatment in cancer.

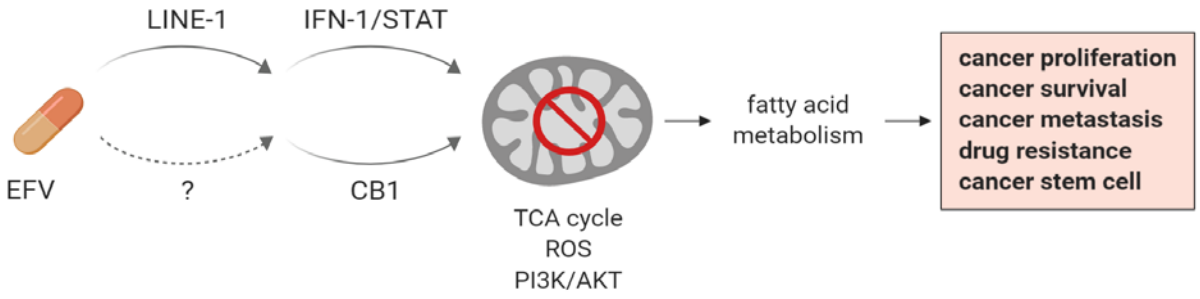


Figure 6-2 predicted connections among various Efavirenz-induced pathways in breast cancers.

6.4 Conclusions

The aim of this project was to show that antiretroviral drugs induce anticancer effects in TNBCs. Antiretroviral drugs can induce cell death, alter cell cycle, and prolong cell proliferation rate in TNBCs. Therefore, antiretroviral drugs are potential anticancer drugs for treating TNBCs. This anticancer process is associated with down regulating cancer fatty acid metabolism, which may be linked to LINE-1 RT inhibition. Furthermore, although we faced many challenges in the study of CSCs, we have used the tumorsphere CSC functional assay to reveal an anti-CSC effect of antiretroviral drugs. Follow-up experiments are necessary to further understand how antiretroviral drugs impact CSC regulation and anticancer activity.

Appendix

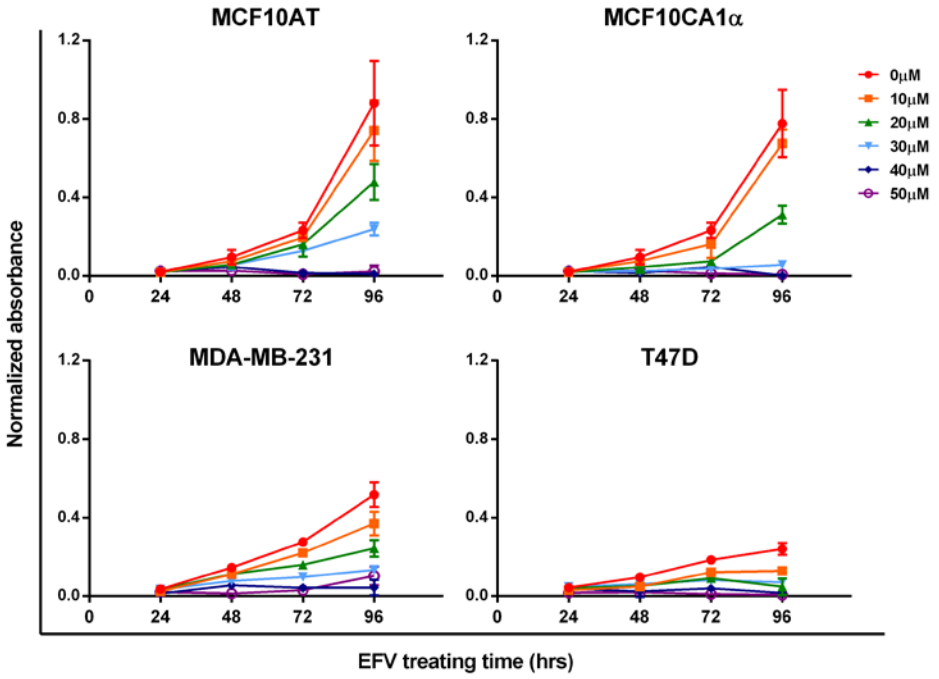


Figure S 3-1 Normalized cell survival curves in different treating conditions

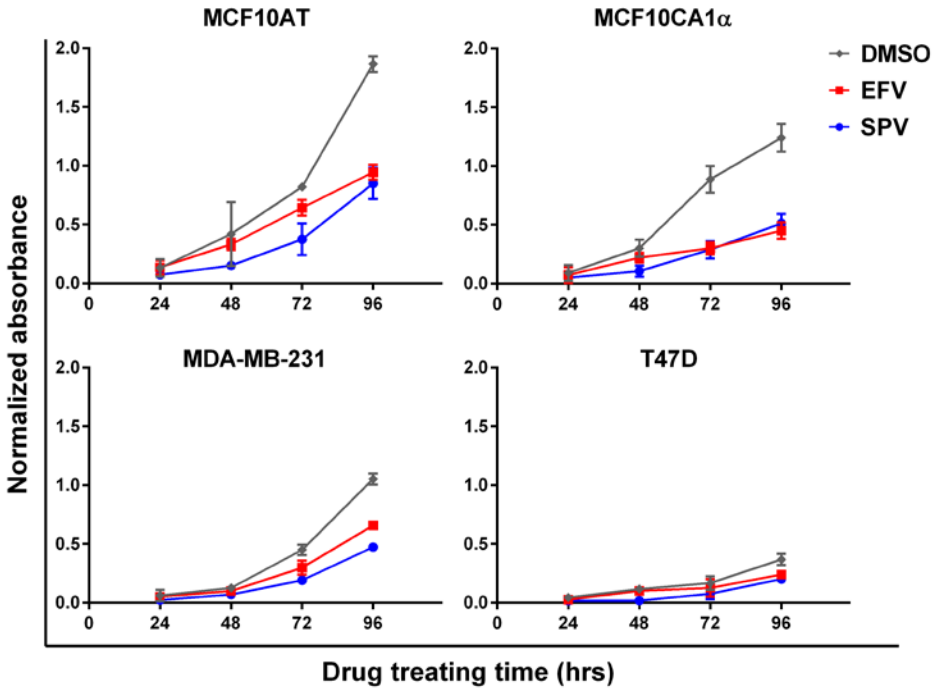


Figure S 3-2 Cell proliferation curves compared antiretroviral drug-treated and control cells for different breast cancer cell lines in different time-points

Table S 3-1 The percentage of cells in different cell phases

		G0/G1 phase	S phase	G2 phase
MCF10A	Control	81.55±9.61%	6.79±3.64%	11.66±6.01%
	EFV	83.22±10.67%	5.77±3.95%	11.01±6.72%
MCF10AT	Control	71.64±7.70%	11.42±2.56%	16.94±6.66%
	EFV	82.22±14.52%	6.02±3.63%	11.76±10.93%
MCF10CA1α	Control	51.48±5.09%	20.39±2.79%	28.13±3.58%
	EFV	63.49±2.00%*	10.81±1.19%**	25.71±1.20%
MDA-MB-231	Control	54.95±1.08%	21.52±3.99%	23.53±5.02%
	EFV	51.91±5.79%	18.81±5.74%	29.28±6.01%
BT-549	Control	58.84±2.29%	19.46±3.46%	21.70±2.17%
	EFV	58.06±3.93%	15.70±0.17%	25.80±4.32%
BT-20	Control	61.87±2.29%	12.61±2.35%	25.52±3.33%
	EFV	53.48±5.01%	9.86±1.51%	36.66±5.05%*
T47D	Control	60.45±6.18%	11.52±1.23%	28.02±5.00%
	EFV	49.51±8.29%	7.65±2.12%	42.84±6.17%*

(* $p<0.05$, ** $p<0.01$)

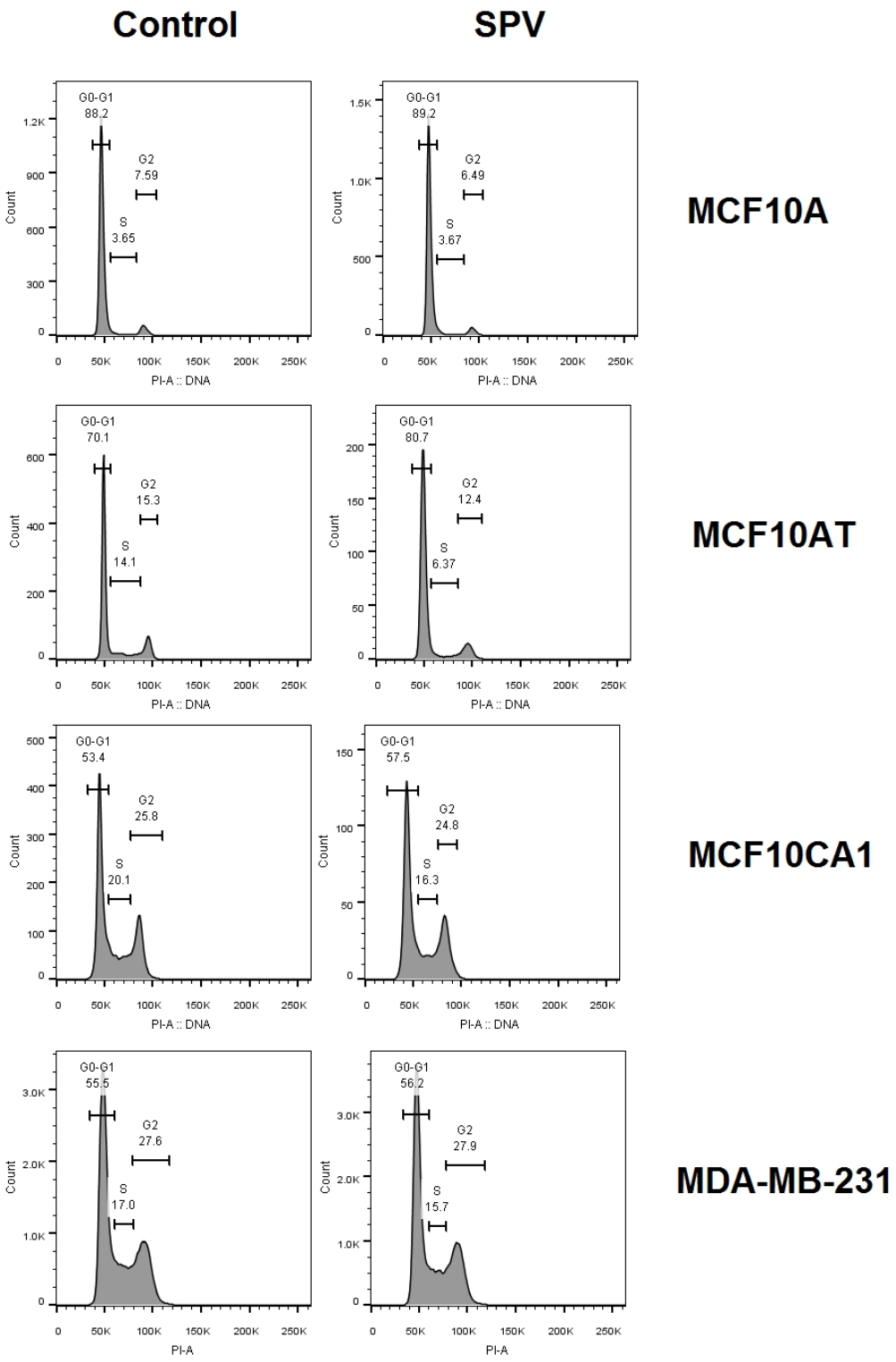


Figure S 3-3 Cell cycle patterns of untreated-control and SPV-treated cells in MCF10A, MCF10AT, MCF10CA1α and MDA-MB-231 cell lines

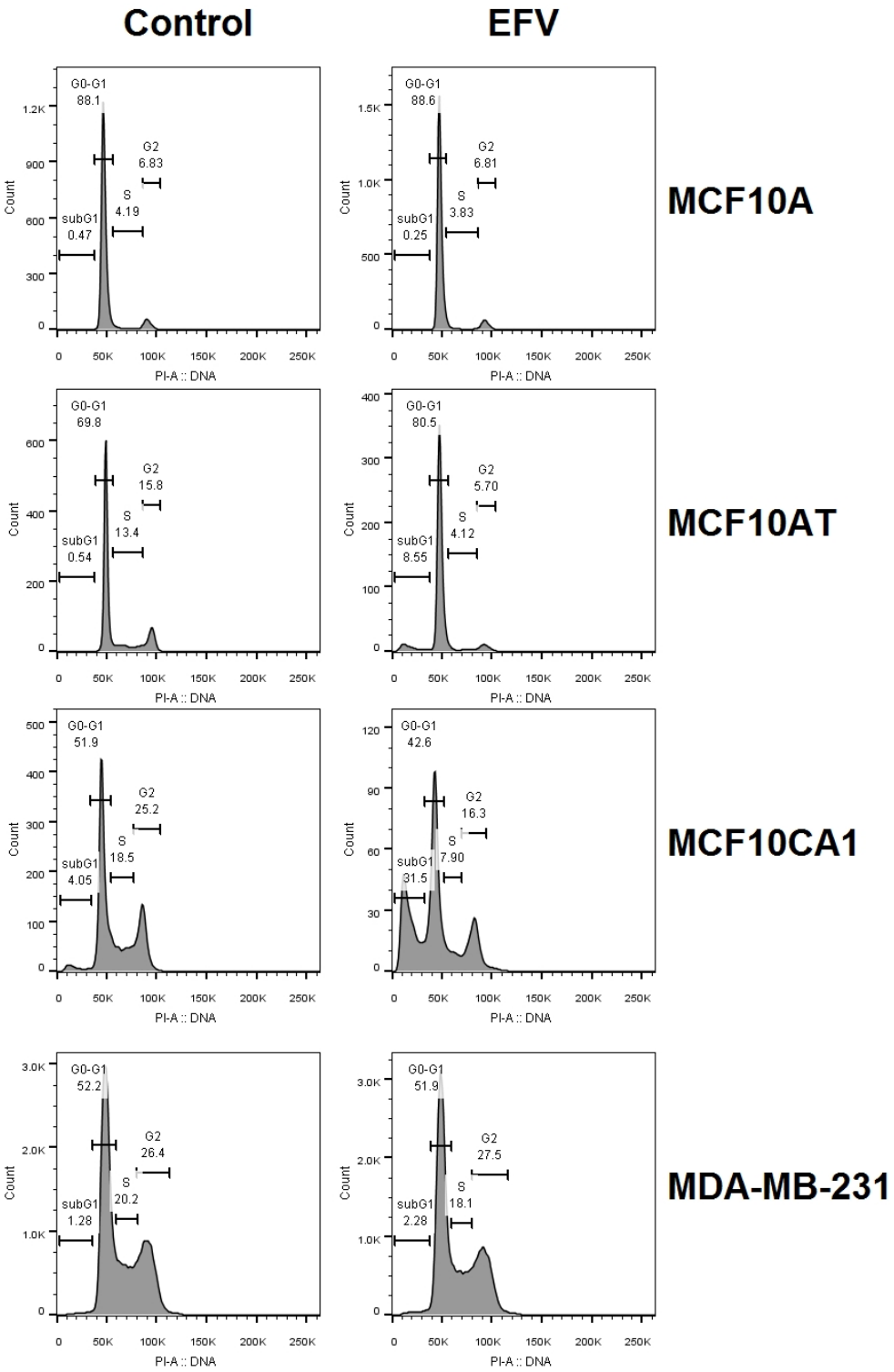


Figure S 3-4 Cell cycle and subG1 patterns of untreated-control and EFV-treated cells in MCF10A, MCF10AT, MCF10CA1α and MDA-MB-231 cell lines

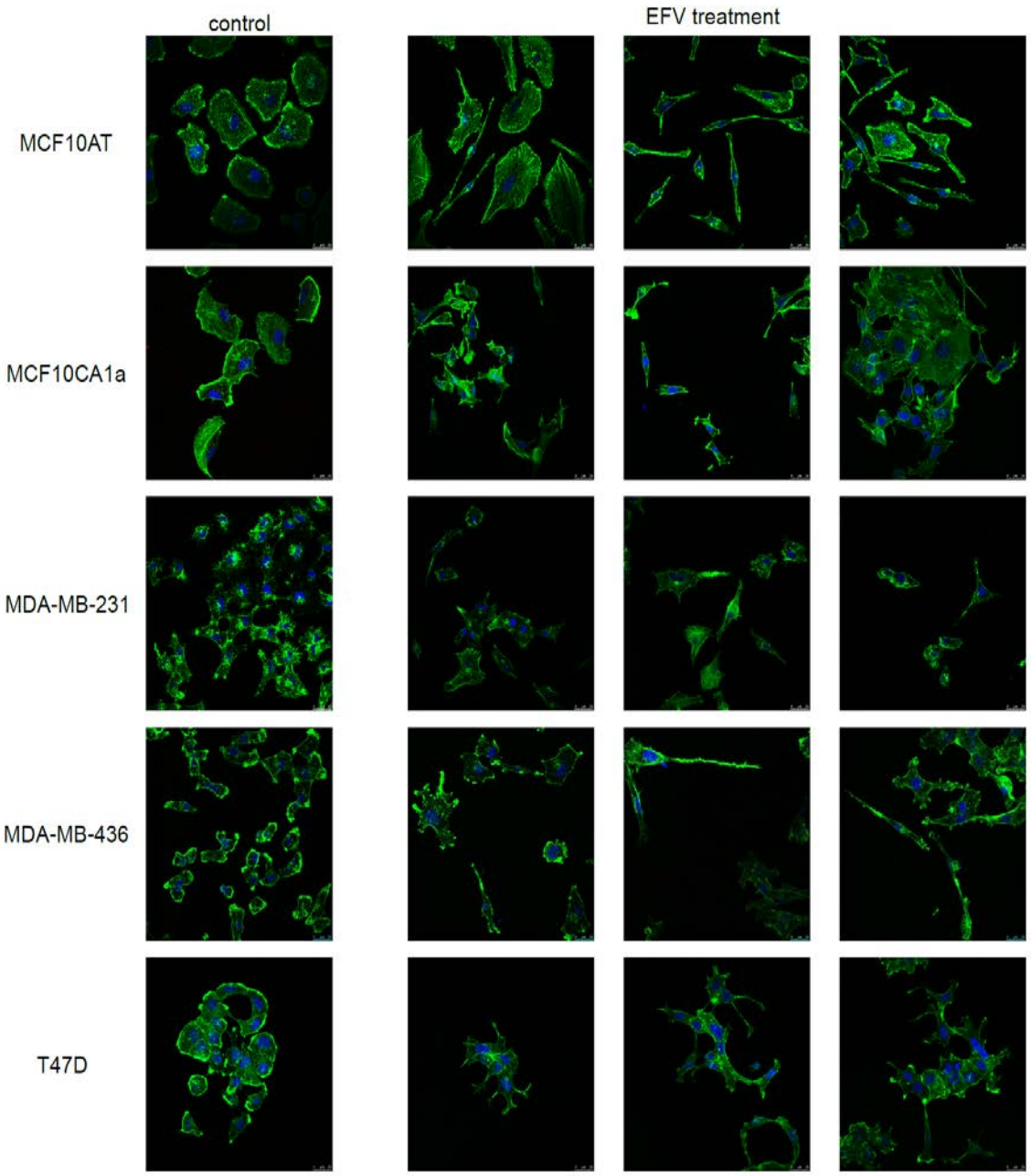


Figure S 3-5 The cell morphologies in EFV-treated and untreated cells
This figure compared the morphologies of EFV-treated and untreated cells in MCF10AT, and MCF10CA1α, MDA-MB-231, MDA-MB-436, and T47D cell lines. Phalloidin (green, for detecting F-actin) and DAPI (blue, for detecting nucleus) was stained.

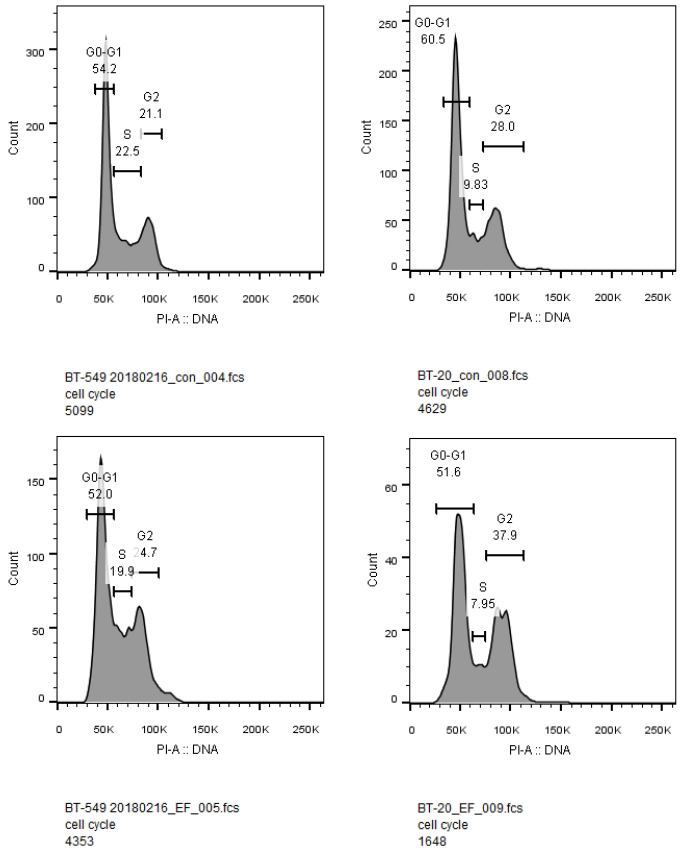


Figure S 3-6 Cell cycle patterns of untreated-control and EFV-treated cells in BT-549 and BT-20 cell lines

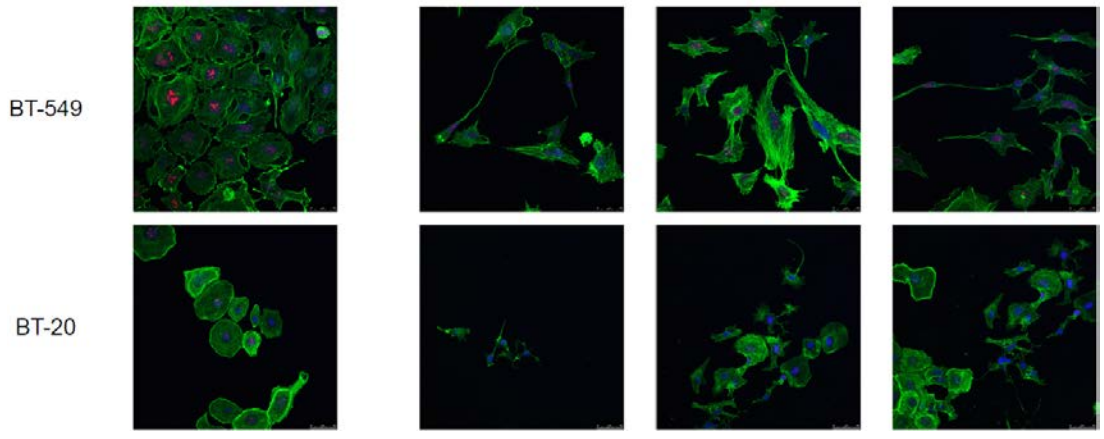


Figure S 3-7 The cell morphologies in EFV-treated and untreated BT-549 and BT-20 cells
This figure compared the morphologies of EFV-treated and untreated cells in BT-549 and BT-20 cell lines. Phalloidin (green, for detecting F-actin), Ki-67 (red, for detecting cell proliferation), and DAPI (blue, for detecting nucleus) were stained.

Table S4-1 EFV down regulation gene set

Gene Name	FC.EF	Gene Name	FC.EF	Gene Name	FC.EF
SCD	-3.79101	RCSD1	-1.63876	MSRB1	-1.44253
ALPP	-3.06442	KLF9	-1.63675	CTSD	-1.43345
PDK4	-2.90026	APOL6	-1.61586	PSME1	-1.42149
ELF3	-2.77384	MXD4	-1.59885	ACOT2	-1.41899
METTL7A	-2.39208	ATOH8	-1.59694	MAN2B2	-1.41881
GBP7	-2.30187	ASAP3	-1.59268	SYT8	-1.41833
SREBF1	-2.29244	PLEKHA7	-1.58839	TGFB2	-1.41469
C3	-2.27458	NCOA7	-1.5818	ACSF2	-1.41449
ACSL5	-2.20747	RHOU	-1.57796	TACSTD2	-1.41328
TMEM135	-2.15772	TFAP2C	-1.57301	HPGD	-1.41211
CYP4B1	-2.15538	IKZF2	-1.57055	SNN	-1.40395
CD14	-2.15152	PRR15	-1.56494	SORT1	-1.40333
FBXO32	-2.13288	TIMP3	-1.56241	DHCR24	-1.39932
RARRES1	-2.07921	SCNN1A	-1.54595	C16orf93	-1.39026
CLIC3	-2.07591	C9orf116	-1.54478	IGFBP4	-1.38172
SGK1	-2.06754	FBLN1	-1.54045	LIPA	-1.37985
KRT13	-1.97777	TIMP3	-1.53326	ZNF362	-1.37266
IFIT3	-1.96808	YPEL3	-1.53273	PDCD4	-1.36833
ETNK2	-1.93086	TLR3	-1.52539	KRT80	-1.36829
BCL6	-1.91379	KCNIP3	-1.52531	PPL	-1.36793
ALDH3A1	-1.89056	ACYP2	-1.51934	VGLL4	-1.36762
TNFAIP2	-1.87149	LINC00511	-1.511	TMEM106C	-1.36037
GRAMD2	-1.84152	LPCAT3	-1.50868	SLC44A2	-1.35994
KLK5	-1.8101	VGLL3	-1.50522	PPAP2B	-1.35529
S100P	-1.80977	SRR	-1.50497	MKL2	-1.35396
SUSD2	-1.7896	A4GALT	-1.49958	CDON	-1.35342
C10orf54	-1.78556	CFH	-1.49874	ABHD8	-1.35018
PLK2	-1.7815	PDK2	-1.47652	YS049_HUMAN	-1.34924
ROM1	-1.75078	FADS1	-1.47548	PSMB8	-1.3482
FASN	-1.73691	CAPS	-1.47507	DDR1	-1.34556
FAM43A	-1.73266	PSCA	-1.47275	SCPEP1	-1.34353
PSMB9	-1.72357	PTPRS	-1.47252	PTPLB	-1.34318
PLA2G16	-1.71928	SLPI	-1.47209	FAM214A	-1.34173
MAP3K8	-1.71266	LPIN1	-1.47064	MIF4GD	-1.33976
NEBL	-1.71162	RAC3	-1.46824	ACACA	-1.33821
ACSL3	-1.69856	SELENBP1	-1.46642	KCNK5	-1.3344
HIST1H3E	-1.67543	SYTL2	-1.46226	C11orf71	-1.33136
NTN4	-1.6714	AC093323.3	-1.45957	RCAN1	-1.33058
PLAC8	-1.66569	PIK3R3	-1.45636	DUSP23	-1.33049
PRSS23	-1.66283	BCL3	-1.45603	LATS2	-1.3235
TSC22D3	-1.65614	VAV3	-1.45298	FXYD3	-1.32013
GLUL	-1.65016	FOXO1	-1.4524	MGAT5B	-1.31981
HERC6	-1.65015	C18orf56	-1.44833	CHMP4B	-1.31691

Table S4-2 continue

Gene Name	FC.EF	Gene Name	FC.EF	Gene Name	FC.EF
PLEKHA2	-1.31666	ASAH1	-1.27268	CTAGE5	-1.24209
DBI	-1.31366	CLN5	-1.27167	DNAJC15	-1.24204
FBXO17	-1.31164	TOB1	-1.27093	ZFYVE28	-1.24193
SRD5A1	-1.31083	TSPYL2	-1.27006	DAB2IP	-1.24078
TP53	-1.30911	CXXC5	-1.26741	PPTC7	-1.24072
CAT	-1.3087	ST20	-1.26684	GPX2	-1.24053
RPP25L	-1.30777	UBL7	-1.26675	NDRG3	-1.23791
PROS1	-1.30758	MADD	-1.26645	MSX2	-1.23779
CDKN2D	-1.30699	NPEPPS	-1.26631	TMEM160	-1.23575
ANAPC16	-1.30322	FAIM	-1.26598	SEMA3F	-1.23573
ARHGEF37	-1.30192	SDPR	-1.26424	PARD6B	-1.2355
SLC25A23	-1.30105	LGALS3	-1.26387	TAF4B	-1.23501
ITFG3	-1.29977	CDC42BPG	-1.2635	DNASE2	-1.23388
ING3	-1.29914	KRT8	-1.25956	ADAT3	-1.22786
OBSL1	-1.29853	TMEM87A	-1.25851	FZD1	-1.22705
KRCC1	-1.29843	ZNF318	-1.25779	BCL7B	-1.2258
COX14	-1.29573	PCCA	-1.2575	MID1IP1	-1.22465
FLYWCH2	-1.2957	CYB5D2	-1.25671	RABL5	-1.22344
DAG1	-1.29469	DHX40	-1.25655	PLEKHA2	-1.21899
MR1	-1.29308	CCDC64	-1.25577	MAML2	-1.21823
FOSL2	-1.28909	ROGDI	-1.2552	VAMP8	-1.21815
DRAM1	-1.28692	ETFDH	-1.25457	PTPRK	-1.21686
PNPLA2	-1.28648	USB1	-1.25352	LAPTM4A	-1.2168
DAPK1	-1.28564	HINT2	-1.25199	IL1R1	-1.21429
GAA	-1.28482	CRAT	-1.25146	MBTPS2	-1.2123
BTN3A2	-1.28264	FRK	-1.25133	STS	-1.20988
ERV3-1	-1.28134	BCL9	-1.25056	IL6ST	-1.20881
NRDE2	-1.28121	FUNDC2	-1.24984	TMEM59	-1.20727
CHP1	-1.28043	NUMA1	-1.24887	LUC7L2	-1.20686
RBPMS	-1.27961	C3orf37	-1.24871	PPM1L	-1.20605
SETMAR	-1.27953	PIP5K1C	-1.2474	TLK1	-1.20531
SLC2A6	-1.27906	ALDH9A1	-1.2454	GCNT2	-1.20367
TSPAN9	-1.27881	MARCH02	-1.24465	PKM	-1.20353
B4GALT1	-1.27743	WWP1	-1.24426	PSMB3	-1.20121
IFITM3	-1.27524	HDAC11	-1.24413	TRAPPC13	-1.20115
GCNT1	-1.27455	ADAM12	-1.24271	GIPC1	-1.20063

(FC.EF: fold-change after EFV treatment)

Table S4-2 SPV down regulation gene set

Gene Name	FC.SP	Gene Name	FC.SP	Gene Name	FC.SP
SCD	-3.6878	PDK4	-1.78511	ATOH8	-1.52911
IFIT3	-2.98949	TLR3	-1.77924	NTN4	-1.5276
ELF3	-2.77005	FAM43A	-1.73073	CTRL	-1.52536
ALPP	-2.74202	TNFSF10	-1.73011	GPR1	-1.52397
RARRES1	-2.60428	MAP2	-1.72985	GLUL	-1.52288
IGFL1	-2.59129	FASN	-1.72564	LIPA	-1.52028
KLK10	-2.57056	C10orf54	-1.72038	IFI35	-1.51973
CD14	-2.48819	APOL6	-1.71969	C9orf3	-1.51958
IFIT2	-2.44074	PRSS23	-1.71253	IGFBP7	-1.5117
GBP7	-2.42212	IFITM2	-1.71252	TSPAN15	-1.5062
METTL7A	-2.42009	PADI2	-1.71227	CCDC64	-1.50211
ISG15	-2.36524	MAP3K8	-1.70926	PARP9	-1.49909
SGK1	-2.36219	PIK3R3	-1.70897	AMOTL2	-1.49801
IFIT1	-2.33476	S100P	-1.69768	FBLN1	-1.49643
TMEM135	-2.32002	ETNK2	-1.69524	PSMB8	-1.49617
ACSL5	-2.31741	SMAD6	-1.68708	TACSTD2	-1.49471
IFI44	-2.26114	IFIH1	-1.66837	NMI	-1.49395
GRAMD2	-2.1697	PRR15	-1.66735	SEPP1	-1.49162
SREBF1	-2.14993	SYTL2	-1.6652	SLC29A3	-1.48672
FBXO32	-2.11818	PLA2G16	-1.65845	GPR110	-1.48293
CYP4B1	-2.06861	TSC22D3	-1.65598	ZNF362	-1.48195
SUSD2	-2.0487	TJP3	-1.65129	VGLL3	-1.48128
PRR15L	-2.01957	SLC43A2	-1.63461	RAC3	-1.48089
PLK2	-2.01621	UTY	-1.60041	ALDH1A3	-1.4795
BCL6	-1.99203	KRT80	-1.59894	BAMBI	-1.47946
NNMT	-1.97731	USP18	-1.59562	SORT1	-1.47944
ID2	-1.96357	KLK6	-1.59059	PSME2	-1.4794
PSMB9	-1.95852	TFAP2C	-1.58136	A4GALT	-1.47784
KRT13	-1.92367	PSME1	-1.57191	BAG1	-1.46806
RIPK4	-1.90997	PROM2	-1.57031	DHCR24	-1.46721
KLK5	-1.89975	KLF9	-1.56828	PDZK1IP1	-1.46637
EFEMP1	-1.89496	SCNN1A	-1.56822	MMP7	-1.46534
HERC6	-1.8633	WISP2	-1.56584	HS6ST1	-1.46415
ALDH3A1	-1.86046	SLPI	-1.55702	FAM214A	-1.46372
AZGP1	-1.84724	SYT8	-1.54988	LIMCH1	-1.46351
ACSL1	-1.83568	LINC00511	-1.54754	CDKN2C	-1.46327
IRF1	-1.83428	PLAC8	-1.54727	FAIM	-1.4623
LCN2	-1.82883	GINS3	-1.54017	ACYP2	-1.46153
CSF1	-1.82685	RCSD1	-1.53784	IFNGR1	-1.46099
NCOA7	-1.82412	PLLP	-1.53776	LRRK1	-1.45507
BCL3	-1.8088	CTGF	-1.53734	ACAT2	-1.45303
TNFAIP2	-1.79212	KRT15	-1.53155	MOB3B	-1.44954
GBP1	-1.78716	NEBL	-1.53139	FADS2	-1.44379

Table S4-2 continue

Gene Name	FC.SP	Gene Name	FC.SP	Gene Name	FC.SP
TSKU	-1.44323	DCN	-1.36781	PDK2	-1.31152
IFITM3	-1.43886	SRSF6	-1.36504	TRAFD1	-1.3104
FXYD3	-1.43408	ING3	-1.36486	TRAF4	-1.3102
TGFBR3	-1.43403	FRAT2	-1.36241	PTPRS	-1.30952
S100A9	-1.43287	ARHGEF37	-1.36052	BST2	-1.30857
MAN2B2	-1.43215	PPAP2B	-1.35692	ACSF2	-1.30819
SMAD1	-1.43062	ASB9	-1.35688	ZC3HAV1	-1.3069
OGFRL1	-1.42692	KCNIP3	-1.3518	ITGAE	-1.30627
ACSL3	-1.42534	SLC25A23	-1.35066	HIST1H4J	-1.30485
OGFRL1	-1.42437	TXNDC16	-1.35025	ARSD	-1.30387
C10orf112	-1.42248	PLEKHA7	-1.34879	TGFB2	-1.3025
HLA-B	-1.42159	ARRB1	-1.34775	HIST1H2AC	-1.3021
MARCKSL1	-1.419	TIMP3	-1.34645	OBSL1	-1.30084
B9D2	-1.41414	LATS2	-1.3421	PTPRK	-1.30057
FRS2	-1.4121	DTX3L	-1.34026	GSDMD	-1.3003
AS3MT	-1.41081	GNRHR2	-1.3394	CTSD	-1.29861
DTWD2	-1.40681	FKBP7	-1.33802	ITGB6	-1.29778
CASP1	-1.40337	GCH1	-1.33709	BCL9	-1.29727
UBE2L6	-1.39911	DUSP23	-1.33627	CDC42BPG	-1.29648
TNFSF13	-1.39869	ATXN7L1	-1.33622	COL2A1	-1.29526
MXD4	-1.39839	PPL	-1.33447	ZNF689	-1.29524
ALG14	-1.39724	TOB1	-1.33437	ASPSCR1	-1.29329
SCPEP1	-1.39137	PLSCR1	-1.33392	SELENBP1	-1.29179
TNS3	-1.39127	DSEL	-1.33244	DTX4	-1.29125
TRIM26	-1.39102	RCAN1	-1.33152	ASAH1	-1.28781
MFAP3L	-1.38965	ZNF628	-1.32781	NSF	-1.28715
TRIM22	-1.38531	NRDE2	-1.32605	HIST4H4	-1.28547
UGT1A3	-1.38479	ARSK	-1.32484	ANXA8L1	-1.28544
CAPS	-1.38358	HIST2H2AA3	-1.32418	TP53	-1.28527
CXCL16	-1.38277	SLC2A8	-1.32262	PLEKHA2	-1.28523
VAV3	-1.38128	SETMAR	-1.32259	HLA-F	-1.28487
FLYWCH2	-1.38004	HIST2H2AC	-1.32235	LANCL2	-1.28473
TP53I3	-1.37954	CFI	-1.3217	WWP1	-1.28448
C9orf116	-1.37946	GLB1L2	-1.32143	SLC25A25	-1.2837
COX14	-1.37711	ST20	-1.32047	DDX58	-1.28356
ACACA	-1.37692	PPCDC	-1.31999	WWC1	-1.28308
PRRG4	-1.37665	FOXO1	-1.31998	CYP2U1	-1.28162
SLC44A2	-1.37658	TUSC1	-1.31738	MSRB1	-1.28039
MRPL54	-1.3751	FAM98B	-1.31705	DBI	-1.28011
PTPLB	-1.37442	DRAM1	-1.31437	ANAPC15	-1.27982
ZBTB22	-1.37285	C21orf2	-1.31347	LYRM2	-1.279
SNN	-1.37064	USB1	-1.31292	STS	-1.27789
CREG1	-1.36792	SLC2A6	-1.3126	GLA	-1.27766

Table S4-2 continue

Gene Name	FC.SP	Gene Name	FC.SP	Gene Name	FC.SP
MID1IP1	-1.27689	LIMK2	-1.24578	TRIM29	-1.22188
KEAP1	-1.27668	B2M	-1.24553	KAT5	-1.22176
DAG1	-1.27504	PROS1	-1.24519	C3orf37	-1.22151
HIST2H2BE	-1.27363	ARHGDIB	-1.24482	ACLY	-1.22103
DHRS4	-1.27226	TMCO4	-1.24463	LAPTM4A	-1.21989
HINT2	-1.27184	C5orf38	-1.24427	IRAK1BP1	-1.2187
ASAP3	-1.26953	USP14	-1.24248	PPM1D	-1.21804
RBPMS	-1.26908	ALDH9A1	-1.24236	C6orf106	-1.21678
CHMP4B	-1.26856	NCBP2	-1.24171	KRT5	-1.21666
KRT8	-1.2683	ECI1	-1.24112	IL1R1	-1.21638
LGALS3	-1.26809	ACOT2	-1.24106	C14orf142	-1.21534
MRPL49	-1.26719	HIST1H2BK	-1.24051	DAB2IP	-1.21466
ATXN10	-1.2655	GBA	-1.2404	PIP5K1C	-1.21357
CSRP2BP	-1.2646	MICB	-1.24027	BTN3A2	-1.21136
SDPR	-1.263	B4GALT1	-1.24018	ARSD	-1.21076
C4orf46	-1.25975	CXXC5	-1.23755	VPS26A	-1.21036
ICK	-1.25937	UBL7	-1.23729	IREB2	-1.20863
LPCAT3	-1.25935	FBXO17	-1.23728	BLOC1S1	-1.20832
SIPA1L3	-1.25915	YS049_HUMAN	-1.237	CTD-2228K2.5	-1.20823
FZD1	-1.25864	PNPLA2	-1.23624	CHST14	-1.20812
FADS1	-1.25845	GPRC5C	-1.23615	SYNC	-1.20797
MGMT	-1.25741	LGALS3BP	-1.23614	EID2	-1.20728
TCF4	-1.25719	CARD6	-1.23513	DDX5	-1.20707
HADH	-1.25688	TSPYL2	-1.23385	IL6ST	-1.20643
CKMT1B	-1.25671	SEMA3C	-1.23334	LINC00610	-1.20583
RFX5	-1.25655	ROGDI	-1.23332	MBTPS2	-1.20558
GOLM1	-1.25456	RILPL2	-1.2324	PKIG	-1.20529
TFDP2	-1.25274	GRWD1	-1.23193	RAET1G	-1.20527
FOXN3	-1.25233	GCNT2	-1.23059	CYB5D2	-1.20487
CNP	-1.25155	FRK	-1.23051	MBIP	-1.20428
CTAGE5	-1.25038	EMC10	-1.23003	RCN1	-1.20404
DPYSL2	-1.2492	ETFDH	-1.22911	PARVA	-1.20386
RHOU	-1.24844	TTC39C	-1.22815	SCP2	-1.20327
TDG	-1.24835	RBBP6	-1.22655	FAM171B	-1.20289
MYEOV2	-1.24813	SH3BGRL	-1.22651	TMEM87A	-1.20188
CCDC94	-1.24737	BID	-1.22638	ADNP	-1.2013
SP2	-1.24728	TPRA1	-1.2258	PKM	-1.20125
SOCS5	-1.24725	TOR1A	-1.22455	SSNA1	-1.20118
CHP1	-1.24656	STRA13	-1.22406	ZNF16	-1.20035
RCC2	-1.24625	FAM185A	-1.22281		
THAP11	-1.24613	LIN52	-1.22254		

(FC.SP: fold-change after SPV treatment)

Table S4-3 EFV up regulation gene set

Gene Name	FC.EF	Gene Name	FC.EF	Gene Name	FC.EF
HMGA2	2.203766	FSCN1	1.486728	NOM1	1.347823
SH2D5	2.153474	OAF	1.472202	FUOM	1.347748
LINC00269	2.014562	FERMT1	1.46415	OCIAD2	1.347692
SMTN	1.980879	ARNTL2	1.460193	GFPT2	1.34698
KDELR3	1.899312	FGFR1	1.459194	HK2	1.3465
PAQR5	1.878873	AC016831.7	1.459101	ICAM3	1.344682
RP11-77I22.3	1.84185	PLEC	1.456763	PLA2G4B	1.342894
HS3ST3A1	1.821999	SEL1L3	1.451235	OSBPL6	1.342822
AXL	1.820712	ERCC1	1.437069	MPP3	1.342511
PLEK2	1.817137	RIN1	1.425565	KIAA0922	1.341844
ARG2	1.811428	TMEM161B-AS1	1.425068	S100A16	1.338613
VIM	1.785889	KLHL17	1.424825	SLC39A13	1.33779
FOXQ1	1.770345	FTL	1.422334	TRMT6	1.335205
FOSL1	1.760571	SMURF2	1.421686	YIPF2	1.335065
LETM2	1.736731	EFNB1	1.416984	LYAR	1.333862
PROCR	1.729174	POFUT2	1.414175	DRAP1	1.333558
SLCO4A1	1.729052	TMEM39A	1.409175	SLC4A3	1.33304
ARHGAP22	1.722266	MLTK	1.404412	MYO10	1.332926
DOCK4	1.702998	OBSCN	1.403007	YS049_HUMAN	1.33207
AC016831.7	1.671331	NAB2	1.397623	MLK4	1.330144
SEC24D	1.667706	IKBIP	1.392482	GCN1L1	1.327157
S100A2	1.653627	ATG16L2	1.391393	ZRSR2	1.326972
STEAP1	1.644226	CCDC112	1.388897	MLKL	1.326049
SLC20A1	1.624318	PKP2	1.388386	KNTC1	1.325738
SDC1	1.592091	ARF4	1.387102	GLS	1.322638
IL15RA	1.590758	COG6	1.380512	SRGAP1	1.321252
RAP1GAP2	1.589534	YS049_HUMAN	1.377235	ATP13A3	1.320405
HYI	1.586216	BCAN	1.375952	L3MBTL3	1.319775
C6orf141	1.584748	SLC29A1	1.375546	TBC1D25	1.314103
HMGA1	1.58358	STON2	1.372851	CALU	1.312283
FLRT2	1.583218	FGFRL1	1.372219	SSR3	1.311618
PID1	1.573859	ITPR3	1.372131	IGF1R	1.310546
DUSP7	1.566578	MAPRE3	1.368488	POLR3D	1.308272
DYRK3	1.547145	DVL1	1.364649	TSC1	1.305009
UAP1	1.539652	FAM135A	1.362991	RMI2	1.303703
ODC1	1.527796	SLC4A7	1.360676	CD274	1.302276
TMEM194B	1.524812	CEP170	1.360516	FBXO22	1.30004
ZBTB14	1.517357	SEC24A	1.359369	GOLGA2	1.297534
AC016831.7	1.511181	B4GALT6	1.355342	ZCCHC6	1.297078
POLR3G	1.505405	SKA1	1.355151	SNAP23	1.297004
HES7	1.503565	MAPKAPK3	1.352218	OVOL2	1.29664
E2F5	1.497455	HES2	1.351096	PFDN2	1.296482
ECM1	1.492682	CD44	1.349677	NAPRT1	1.294557

Table S4-3 continue

Gene Name	FC.EF	Gene Name	FC.EF	Gene Name	FC.EF
MED8	1.29387	TTL	1.257991	RRBP1	1.228391
MSTO1	1.292491	ATM	1.257887	ZNF83	1.228033
GM2A	1.291962	PIBF1	1.25779	GALE	1.227444
FAM83G	1.291807	DICER1	1.257408	CERS2	1.22611
EPHB2	1.291595	KCTD21	1.256589	SACS	1.223408
TTC14	1.291463	ZCCHC7	1.256488	SLC9A3R2	1.223368
YS039_HUMAN	1.289468	ABCC4	1.25638	CMTM7	1.222058
DEPDC1	1.287548	PJA1	1.25629	GAS2L3	1.221804
GINS2	1.286417	LEPREL4	1.254293	ZNF562	1.221443
STIM2	1.28432	POPDC3	1.253956	RAB21	1.218739
INPP4B	1.283642	GREB1L	1.253735	ATXN2L	1.218199
WDR27	1.283417	EBLN1	1.253607	LAD1	1.216748
GFPT1	1.282795	BDKRB2	1.252392	PPIL4	1.216377
RP11-640M9.2	1.282666	MAP7D3	1.252236	AASS	1.216155
PIK3CA	1.280802	FXYD5	1.252208	DPY19L4	1.215643
RP11-319G6.1	1.280419	CAMSAP2	1.252085	SCLY	1.215479
ALAS1	1.27808	MMP1	1.250558	WDR35	1.214325
CORO1C	1.277469	CBWD6	1.250514	DNMT3A	1.212779
FKBP10	1.275518	EIF4EBP1	1.249676	RALB	1.212321
ARFGAP1	1.275248	COG3	1.249182	SIRT7	1.211955
CTSC	1.273337	FARSB	1.248906	SURF4	1.211857
PAQR3	1.272657	FANCB	1.248838	COPA	1.2095
AGO2	1.271307	SOWAHC	1.246658	P4HB	1.209468
PLXNA2	1.270281	WDR24	1.246377	DGKE	1.20937
BCDIN3D	1.26979	ELP2	1.246321	FAM98A	1.209233
MACROD1	1.268196	CNTROB	1.244822	MAP3K9	1.209142
DUS4L	1.267882	STRIP2	1.241882	COPB2	1.207597
YIF1A	1.267764	Shc1	1.240892	KIAA0020	1.207252
LIN1_NYCCO	1.267615	SLC15A4	1.240253	UTP6	1.207213
FRAS1	1.266537	SH3D21	1.239871	PTPDC1	1.206412
CCDC144B	1.266015	RTTN	1.23973	MON2	1.205943
ZDHHC8	1.265395	GTPBP10	1.238081	RGS10	1.205577
QSOX2	1.264585	SLC30A5	1.237782	MGEA5	1.205509
LARP1B	1.263485	ZNF280C	1.237242	ZNF37A	1.204914
PPAPDC1B	1.263242	SETX	1.236806	DNAH14	1.204586
MAPKBP1	1.26304	LRRC8E	1.23658	ZNF7	1.203814
ACBD3	1.262892	MAP7D1	1.234634	UBA6	1.203583
KCTD12	1.262795	ERCC5	1.233921	UFL1	1.203487
AC004980.7	1.262483	CDK17	1.2338	YG039_HUMAN	1.203381
BRI3	1.26089	SHARPIN	1.232926	NCS1	1.203188
LRP8	1.259807	TTC3	1.23093	STEAP3	1.203123
UHRF2	1.259199	ATG12	1.230763	DIS3	1.203102
ZNF860	1.259025	RBM28	1.2303	ZNF841	1.202662
MDM4	1.258134	NXPH4	1.228813	FBXO45	1.202094

Table S4-4 SPV up regulaiton gene set

Gene Name	FC.SP	Gene Name	FC.SP	Gene Name	FC.SP
ANGPTL4	2.77369	SH2D5	1.663312	MXRA8	1.491749
CAPRIN2	2.428073	SEC24D	1.661752	UBR4	1.486684
RPSAP52	2.33773	SCNN1D	1.657144	UACA	1.485385
SERPINE1	2.300503	PROCR	1.657094	EMP3	1.484684
VIM	2.255324	RP11-290L1.3	1.641563	LRP1	1.48436
HMGA2	2.214457	LETM2	1.635522	DENND4B	1.483972
PAQR5	2.203174	UTY	1.616979	FLRT2	1.483867
IL13RA2	2.164794	WDR27	1.611014	RIN1	1.481809
C9orf85	2.162859	ENO3	1.602986	BMS1	1.476621
CREB5	2.094372	AC016831.7	1.601013	AHSA2	1.472717
SMTN	2.072298	ARHGAP22	1.59936	GPR153	1.471915
RP11-77I22.3	2.032499	DZIP1L	1.59334	YS049_HUMAN	1.468911
HS3ST3A1	2.022767	PID1	1.591747	DBN1	1.46542
RTEL1	2.014424	GG6L5_HUMAN	1.584375	LARP6	1.463662
SLCO4A1	1.999081	MSH5	1.566028	C11orf68	1.46196
C6orf141	1.994882	DGKA	1.562121	DUSP7	1.45896
FOXQ1	1.977663	SDC1	1.561437	LRRC27	1.458848
AXL	1.969879	PTK2B	1.557532	OAF	1.458728
DUSP6	1.962258	ITGA3	1.552132	POLG2	1.456494
MTMR11	1.906216	PLEC	1.549832	ANXA5	1.451464
AGPAT9	1.882447	HYI	1.548615	RAP1GAP2	1.446288
FOSL1	1.878668	ENTPD2	1.547297	FSCN1	1.439189
SEMA3B	1.856047	NRG1	1.546116	SPTB	1.433211
UBE3D	1.819088	RP11-253M7.1	1.539064	SFN	1.431464
SERPINE2	1.791818	CCDC88A	1.536684	UBXN11	1.429667
HES7	1.783064	EFNB1	1.535182	GPSM1	1.427929
FAM196B	1.771963	PLXNA3	1.53406	POLM	1.425913
NLRP1	1.768953	YS049_HUMAN	1.53038	MBOAT2	1.425401
PLEK2	1.765578	PLCH2	1.529903	DYRK3	1.423515
VGF	1.763073	FMNL2	1.528466	GFPT2	1.42227
SLC20A1	1.754769	ARHGEF2	1.524077	FTL	1.421658
S100A2	1.740106	FGFR1	1.517815	ABLIM3	1.419424
KDELR3	1.738288	DOCK4	1.517081	C7orf41	1.417534
PLEKHG5	1.725315	TMEM194B	1.510499	PI4KAP2	1.413476
AHNAK2	1.714634	AC016831.7	1.507974	JRK	1.409942
S100A6	1.712042	ECM1	1.50755	ADAT2	1.409909
HMGA1	1.704836	ZNF485	1.506275	PGM3	1.408842
CDKL1	1.698651	ZNF506	1.499848	CCDC41	1.408263
ATG16L2	1.694391	RAC2	1.499239	AC016831.7	1.40751
AC016831.7	1.68458	DPP3	1.495876	UAP1	1.40726
STEAP1	1.683758	S100A6	1.49576	BCAN	1.407103
C16orf89	1.683289	SLCO4A1	1.494883	DVL1	1.406629
FAM83A	1.681807	TEP1	1.494317	LTBP4	1.405844

Table S4-4 continue

Gene Name	FC.SP	Gene Name	FC.SP	Gene Name	FC.SP
IER2	1.404662	DEPDC1	1.362589	CBWD5	1.32714
LPXN	1.403281	TPD52L2	1.361281	THBS3	1.326592
ENO4	1.400903	C15orf52	1.360169	LINC00269	1.326539
RP11-319G6.1	1.400839	GCN1L1	1.359618	TMEM120A	1.326063
MST1R	1.399882	LIN1_NYCCO	1.3596	ACAD10	1.325066
Gtf2i	1.398289	LIN1_NYCCO	1.359154	SLC4A7	1.324964
CLMP	1.397945	ZNF333	1.357919	TMEM136	1.323455
ERCC1	1.397625	ZNF516	1.357314	KATNAL1	1.323271
SPAG9	1.396749	RP11-640M9.2	1.357198	FAM135A	1.323096
CEP170	1.396561	GOLGA8A	1.357097	FAM83G	1.322955
SH3D21	1.395901	SLC4A3	1.353571	PLXNA2	1.322211
MCTP1	1.39533	ZNF860	1.353334	UTRN	1.321851
IL15RA	1.394598	PTK6	1.353084	PLA2G4B	1.320935
CCDC144B	1.39321	SRGAP1	1.351688	GDI1	1.320748
KLHL17	1.391468	FAM86DP	1.351417	TMEM161B-AS1	1.320602
B4GALT6	1.390726	SDHAP1	1.350897	S100A16	1.320417
C9orf72	1.38869	NPHP3	1.350837	N4BP2	1.318962
CDC25B	1.387614	ACAP3	1.349888	GLS	1.318457
DOCK6	1.387181	MAR09	1.349686	CDRT4	1.318273
APAF1	1.386312	SEL1L3	1.349042	TNRC6B	1.317297
LYST	1.382528	PAX8	1.348035	MAPKBP1	1.317209
ARNTL2	1.380779	DGKQ	1.347359	HEMK1	1.317069
WSB1	1.380566	POFUT2	1.347282	YS039_HUMAN	1.316903
SZT2	1.380501	CCZ1	1.345921	SMURF2	1.316462
Fv1	1.37986	DUSP22	1.344925	ZNF160	1.316066
ITPR3	1.379608	RGS20	1.343286	C16orf89	1.315541
IGF2BP2	1.379079	NDE1	1.342362	GOLGA2	1.315492
ATP8B3	1.378308	HK2	1.34222	ENO4	1.315056
FGFRL1	1.376544	VAT1	1.342065	PPP1R13L	1.314353
C9orf85	1.376017	MPP3	1.340909	TMEM39A	1.314285
NAB2	1.373861	ELMO3	1.340509	C5orf54	1.314208
FOXD1	1.373238	DMTF1	1.339711	PPP2R2C	1.314145
WDR60	1.372969	IQGAP3	1.337453	DUS4L	1.313615
MYO10	1.372665	PVRL1	1.336414	KIF27	1.313429
MNT	1.372274	NOM1	1.3363	NFAT5	1.313277
SEMA4B	1.37098	DUOX1	1.334319	NBPF9	1.313132
DRAP1	1.370899	CCDC112	1.332931	FCHSD1	1.312741
MYADM	1.37039	TTC14	1.332313	FERMT1	1.311019
MLTK	1.36813	MACROD1	1.331366	ANKRD36C	1.310788
MAPKAPK3	1.367452	ZCCHC6	1.3304	AGO2	1.310733
SNX29P2	1.365968	DNMT3A	1.328232	KCNAB2	1.310494
ANKRD36	1.364055	KLC3	1.327846	GLI4	1.309809
ESPL1	1.363538	TLDC1	1.327695	CMTM7	1.309568

Table S4-4 continue

Gene Name	FC.SP	Gene Name	FC.SP	Gene Name	FC.SP
CCBL1	1.30901	PTPDC1	1.284203	MTPN	1.258185
LRP8	1.307025	CD44	1.283369	CD276	1.257955
LINC00704	1.30622	MSTO1	1.282942	CALU	1.256646
FKBP10	1.304481	AMOTL1	1.282537	ASPHD1	1.256227
DMXL2	1.304291	ZNF330	1.282214	MORN1	1.256089
SEC24A	1.304204	ALAS1	1.281455	STRIP2	1.254393
IKBIP	1.303851	OSBPL6	1.281395	ODC1	1.253856
RMI2	1.302999	RIOK3	1.281275	SLFN13	1.253364
NBPF20	1.302606	UHRF2	1.279618	EPHA2	1.253147
E2F5	1.302514	SSR3	1.279338	AP4B1	1.252445
NAPRT1	1.302225	STK32C	1.279304	MAP3K10	1.25167
PAK1	1.301141	FUOM	1.278153	NFRKB	1.250699
DPAGT1	1.301037	LONRF1	1.277813	Shc1	1.250116
LRRFIP1	1.300555	PTGR1	1.277694	LIFR	1.250041
CPNE8	1.300375	MLK4	1.276749	DDHD1	1.249984
PHF7	1.299924	NKTR	1.275888	CERS2	1.249646
DLG4	1.299042	MALAT1	1.274934	CBWD6	1.248852
DDHD2	1.298854	DYNC1H1	1.274726	TTC3	1.24864
DENND3	1.298064	L3MBTL3	1.274369	YEATS2	1.247727
Csnk1e	1.297921	BRI3	1.274272	NBPF24	1.247528
ATP13A3	1.297723	CBLC	1.273868	PAQR3	1.245976
RASA3	1.296506	C1orf132	1.273303	SENP7	1.245506
MDM4	1.295934	RGL3	1.272768	GLYCTK	1.243998
POLR3G	1.295764	MEX3C	1.272689	YG039_HUMAN	1.243829
SLC45A3	1.294726	ZNF431	1.272322	ZCCHC7	1.242925
LRRC37B	1.294163	FBXO22	1.272054	CNTROB	1.242798
CCDC180	1.294027	NT5C	1.271479	PPP5D1	1.242328
NFKBID	1.29375	CAMK2D	1.269319	ANKRD36B	1.242252
ECI2	1.292831	MON2	1.268876	MICAL1	1.242165
C9orf40	1.292207	SLC15A4	1.268524	GMPPA	1.241443
TMEM55A	1.290954	RALB	1.267795	NPIPB5	1.241417
NOC3L	1.290138	COG1	1.267699	KATNBL1	1.240862
CTTNBP2NL	1.289566	GFPT1	1.267283	ATM	1.240315
RTTN	1.288938	CD97	1.267232	USP4	1.239806
YIPF2	1.288812	RFFL	1.267156	SLCO1B3	1.239638
GINS2	1.288613	RASA2	1.267149	PNISR	1.239514
AC004980.7	1.288493	C12orf23	1.266577	AMT	1.238918
CEP290	1.287922	CYP2R1	1.265165	CXorf57	1.237152
GPR161	1.287295	ZNF200	1.264983	FAM157C	1.237131
OBSCN	1.286841	PIBF1	1.262176	MYEOV	1.236623
FRAS1	1.285284	OCIAD2	1.261651	ZNF337	1.236607
CYP26B1	1.284592	ZNF841	1.260813	GALE	1.236446
SIM2	1.284504	HAUS6	1.260064	ASUN	1.236339

Table S4-4 continue

Gene Name	FC.SP	Gene Name	FC.SP	Gene Name	FC.SP
THAP3	1.236292	ZDHHC8	1.225524	LRIG2	1.211889
GPSM2	1.23623	EPHB2	1.224218	ZNF529	1.211425
ZNF550	1.235739	ACADVL	1.223617	LINC00269	1.209953
PJA1	1.235374	ZRSR2	1.223342	SESTD1	1.209208
FAM219A	1.234886	C16orf58	1.222856	SETX	1.209072
MAP3K9	1.23481	ACBD3	1.222784	COG3	1.208857
RRBP1	1.23463	C20orf194	1.222524	COTL1	1.208726
GM2A	1.234194	FAM73A	1.222482	FBXL4	1.208502
RBM26	1.234137	MAN2C1	1.222348	GON4L	1.208195
PPAPDC1B	1.234054	PIP4K2A	1.222317	MYO9B	1.20798
RP11-395B7.7	1.233964	RRNAD1	1.22227	CEP72	1.207678
UVSSA	1.23349	NF1	1.221505	YIF1A	1.20736
PTAR1	1.233271	SHARPIN	1.2214	MPRIP	1.207189
ASPH	1.233185	TRMT6	1.221335	CEP135	1.20677
KCMF1	1.233037	PLEKHA8	1.221237	SMG1	1.205199
LEPREL4	1.23269	CBL	1.22102	RECQL	1.205057
PRRC1	1.232556	ARHGEF35	1.220941	JPH1	1.205054
RNF214	1.232418	CASC5	1.220047	HDDC3	1.204835
TNKS	1.23221	MNT	1.218489	ZNF83	1.204662
CEP120	1.232049	LARP1B	1.218237	AP1S1	1.204553
UBA6	1.231855	SLC9A3R2	1.218201	TAZ	1.204546
SGTB	1.23179	GLUD1	1.217669	TSC1	1.203925
ODF2L	1.231696	CERK	1.217254	MAP7D3	1.203677
TTL	1.231333	ATG12	1.21708	NCOA3	1.203155
UTP6	1.230092	POLE4	1.216742	L3MBTL1	1.20306
GREB1L	1.22986	ODF2	1.216712	UGGT2	1.202528
PFDN2	1.22937	RCOR1	1.215978	RILPL1	1.20244
LMAN1	1.22932	MLH3	1.215939	DGKE	1.202275
AAED1	1.228838	NAT9	1.214909	TAOK3	1.202115
TUBG2	1.228637	UFL1	1.214888	ATP11A	1.202111
SNAP23	1.228627	KIFAP3	1.214678	BTAF1	1.20179
RTCA	1.228067	GTF3C3	1.214499	RFNG	1.201537
SLC39A13	1.227137	GALNT3	1.21404	KIAA1549	1.201203
TRAF5	1.226961	FAT1	1.213614	SCRN1	1.201192
RSBN1	1.226587	NET1	1.213223	TTBK2	1.200982
SIRT7	1.225741	LINC00269	1.213217	FAM126B	1.200614
FAM200B	1.225737	DHRS7	1.21297	RP9	1.200425
HACE1	1.225543	CTSC	1.212947		
PRKAA2	1.225529	FARSB	1.212084		

(FC.SP: fold-change after SPV treatment)

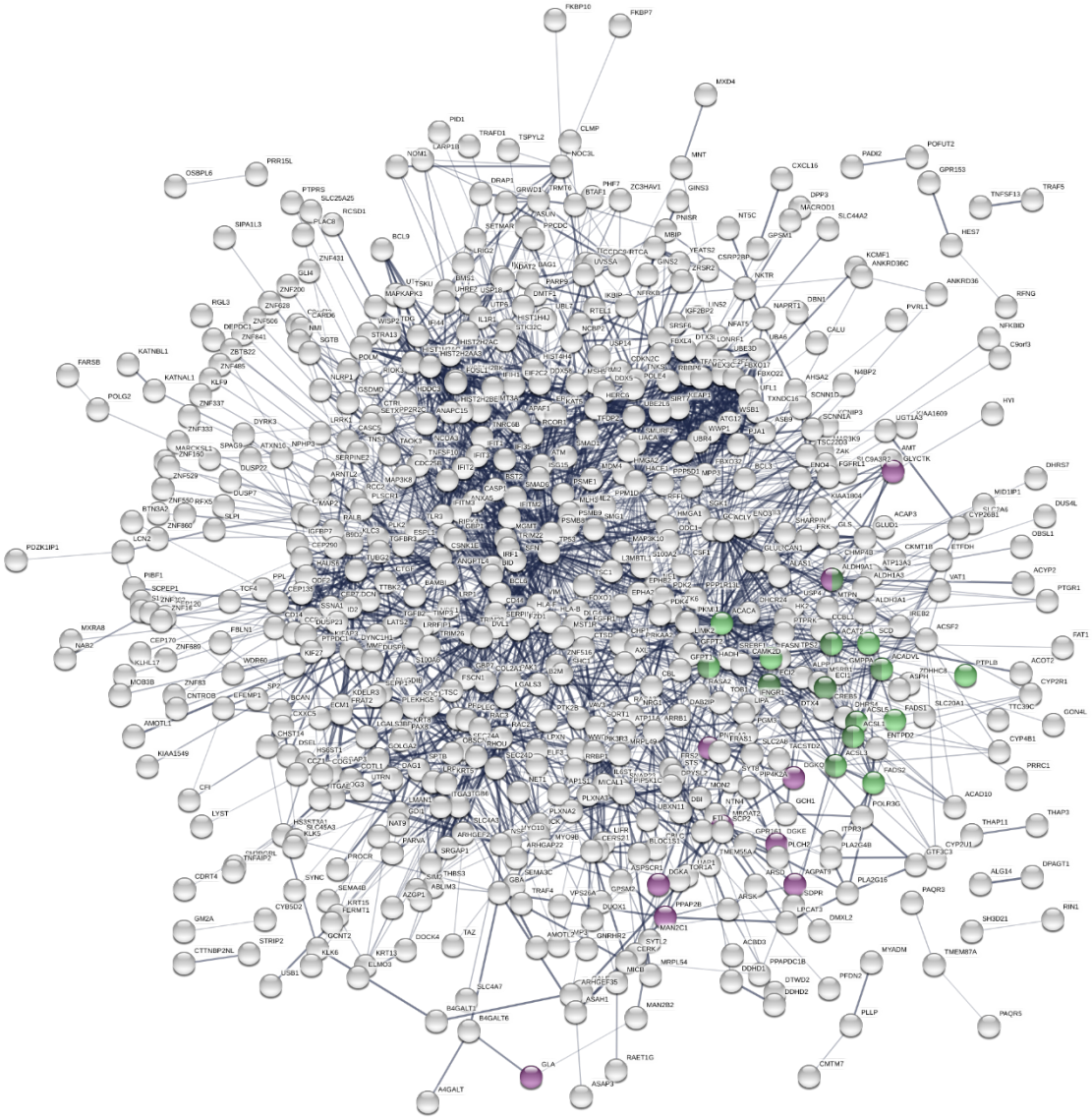
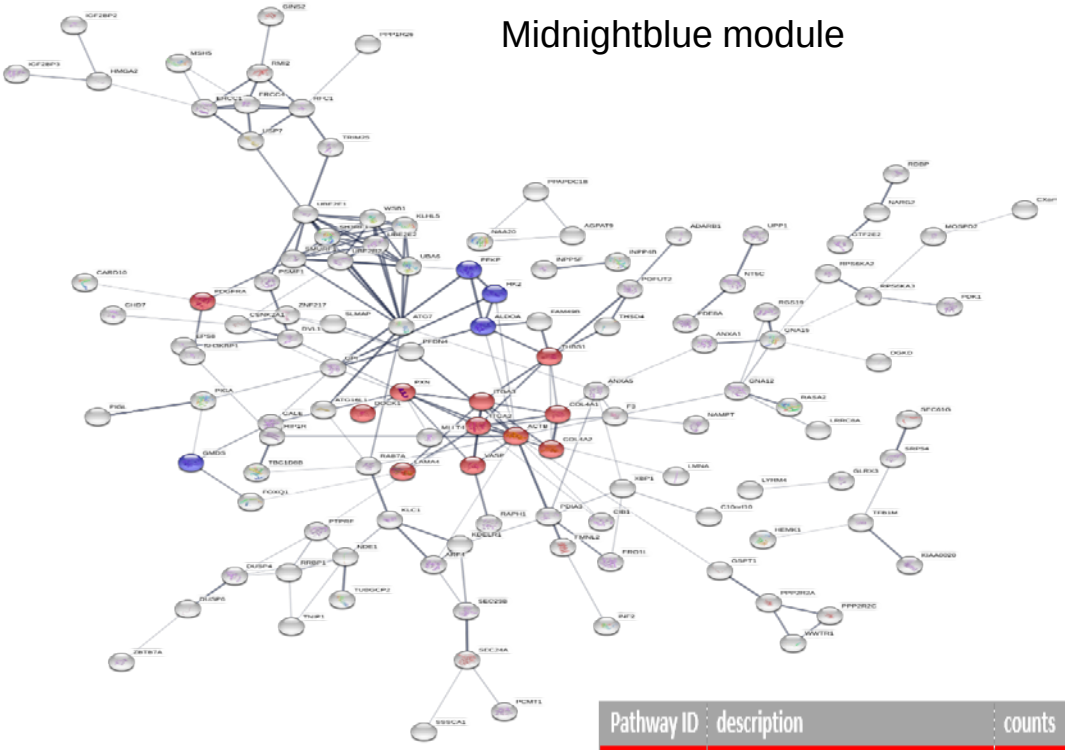


Figure S4-1 The STRING-DB gene network for SPV-treatment induced up regulated and down regulated genes in TNBCs

Table S4-5 Highlighted KEGG pathways in SPV regulated gene sets using STRING-DB

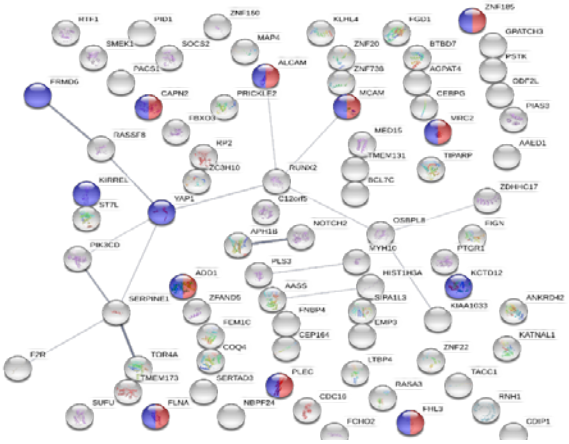
Pathway ID	Pathway description	Count network	*q*-value
01212	**Fatty acid metabolism**	7	9.93×10^{-5}
00071	**Fatty acid degradation**	9	0.00534
00561	**Glycerolipid metabolism**	10	0.0534

Midnightblue module



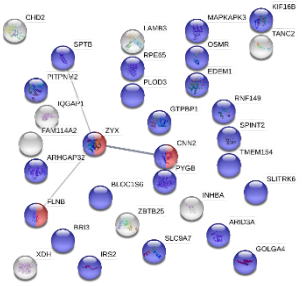
Pathway ID	description	counts	q-value
04510	Focal adhesion	11	0.00152
00051	Fructose and mannose metabolism	4	0.251

Darkorange module



Pathway ID	description	counts	q-value
GO:0005925	Focal adhesion	9	0.00659
GO:0030051	Cell junction	13	0.0465

Orangered4 module



Pathway ID	description	counts	q-value
GO:0001725	Stress fibre	3	0.0281
GO:0016020	membrane	24	0.0281

Figure S4-2 Specific pathways being highlighted in different TNBC-only network analysis modules

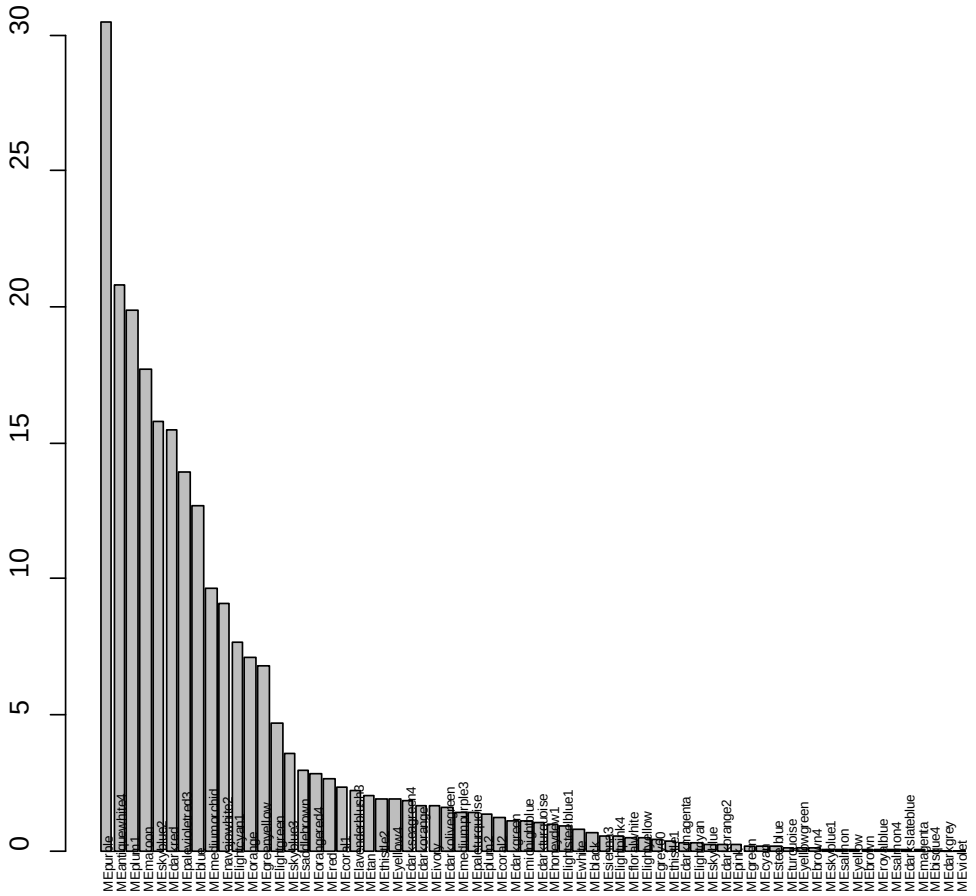
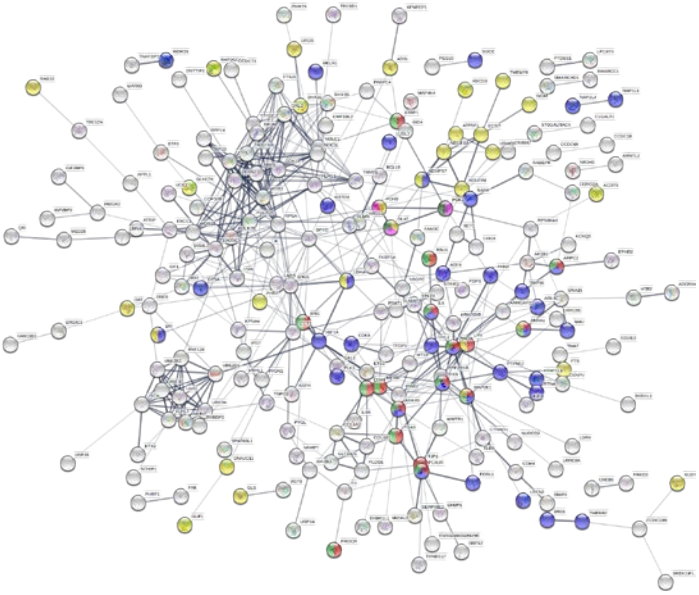


Figure S4-3 Bar Plot of ANOVA *q*-values for module eigengenes in the network analysis (TNBCs+MCF7)

Table S4-6 ANOVA *q*-values for module eigengenes in the network analysis (TNBCs+MCF7)

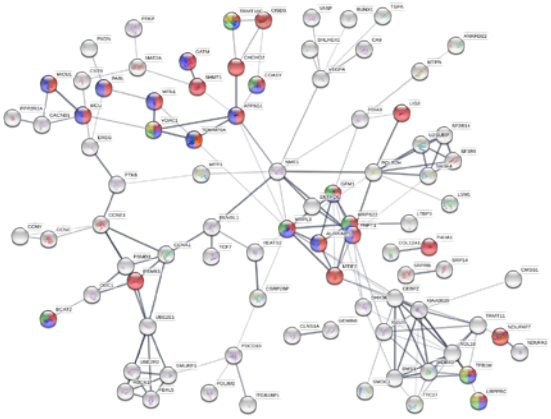
Module name	*p*-value
MEpurple	0.02445472
MEantiquewhite4	0.03491513
MEplum1	0.03491513
MEmaroon	0.03764551
MEskyblue2	0.03764551
MEdarkred	0.03764551
MEpalevioletred3	0.04421653
MEblue	0.04997094

Purple module



Pathway ID	description	counts	*q*-value
GO:0005912	Adherens junction	19	0.00751
GO:0042995	Cell projection	46	0.00904
GO:0005925	Focal adhesion	17	0.00971
GO:0005739	Mitochondrion	45	0.0112
GO:0045254	Pyruvate dehydrogenase complex	3	0.0157
GO:0005967	Mitochondrial pyruvate dehydrogenase complex	2	0.0336

Darkred module



Pathway ID	description	counts	*q*-value
GO:0005739	Mitochondrion	32	2.59×10^{-5}
GO:0044429	Mitochondrial part	21	0.000173
GO:0005759	Mitochondrial matrix	11	0.00613
GO:0042645	Mitochondrial nucleoid	4	0.121

Figure S4-4 Specific pathways being highlighted in different TNBC+MCF7 network analysis modules

Table S4-7 Mitochondria-related genes in network analysis

Gene Name	FC.EF	FC.SP	Gene Name	FC.EF	FC.SP
ABCD3	-1.1386	-1.14812	MFN1	1.17084	1.14744
ACN9	-1.15719	-1.10682	MICU1	-1.03522	-1.07097
ACOT9	1.17204	1.17915	MRPL3	1.05854	-1.00796
ADSL	1.12105	1.04844	MRPS22	-1.00912	1.02331
ANXA1	1.10793	1.1234	MRPS27	-1.01285	-1.01285
ATP5F1	-1.0012	-1.06734	MRS2	1.01629	-1.05724
ATP5G1	1.04382	-1.04646	MTIF2	1.00186	1.00748
AURKAIP1	1.0188	-1.00069	NDUFA8	1.02355	-1.02442
BCAT2	1.09678	1.03948	NDUFAF7	1.09973	1.05501
C10orf10	1.376	1.60306	NDUFS4	1.08402	-1.01554
CAPRIN2	1.84355	2.05798	NDUFS7	-1.01483	-1.07597
CCDC58	1.08177	1.0672	NUDT13	-1.18834	-1.01491
CHCHD2	-1.00637	-1.06191	NUDT6	1.05245	-1.06933
CISD3	1.0866	1.01239	OAT	-1.07609	-1.12787
CLPP	1.05066	-1.02805	OCIAD2	1.17902	1.09682
COA3	-1.0176	-1.02064	P4HA1	1.01299	1.14766
COASY	1.02873	1.00908	PARL	-1.06503	-1.11942
DHX32	-1.01648	-1.03279	PCCB	-1.00389	-1.01368
DLAT	1.00326	-1.04379	PDHB	-1.07431	-1.08912
DNAJC11	1.03761	-1.05997	PDK1	1.08806	1.13691
ECSIT	-1.05042	-1.12539	PHB2	-1.00531	-1.01183
ELK3	1.22842	1.26687	PNPT1	-1.0155	-1.06202
ETFB	-1.11341	-1.03465	PRELID2	1.18676	1.27161
GATM	-1.26724	-1.18385	PSMB3	-1.14617	-1.15775
GCAT	1.10026	1.13098	PTS	1.06382	1.02226
GFM1	-1.02747	-1.06987	RAB32	1.06249	1.00269
GLS	1.27994	1.28803	RAI14	1.05627	1.22371
GLYCTK	1.0942	1.19769	RAP1GDS1	1.11381	1.06734
GUF1	1.08906	1.04166	SFXN4	1.00284	-1.01034
HEMK1	1.23966	1.34036	SHMT1	-1.10275	-1.1961
LDHA	-1.06593	-1.03196	SRI	1.03843	-1.01021
LETM2	1.59791	1.46854	TFB1M	1.20072	1.16305
LIG3	1.10457	1.02883	TMEM70	1.03425	-1.07544
LRPPRC	1.04112	-1.01491	TOMM70A	1.04083	-1.01414
LYN	1.03942	-1.02029	TRAK1	1.10276	1.17932
LYRM4	1.16338	1.04836	TRMT10C	1.06115	-1.0453
MAP2K1	1.19618	1.1732	UROS	1.03531	-1.02063
MCU	1.03248	-1.01273	USMG5	1.06013	1.02589
MFF	1.07091	1.02748	VDAC1	1.04226	1.0086

(FC.EF: fold-change after EFV treatment; FC.SP: fold-change after SPV treatment)

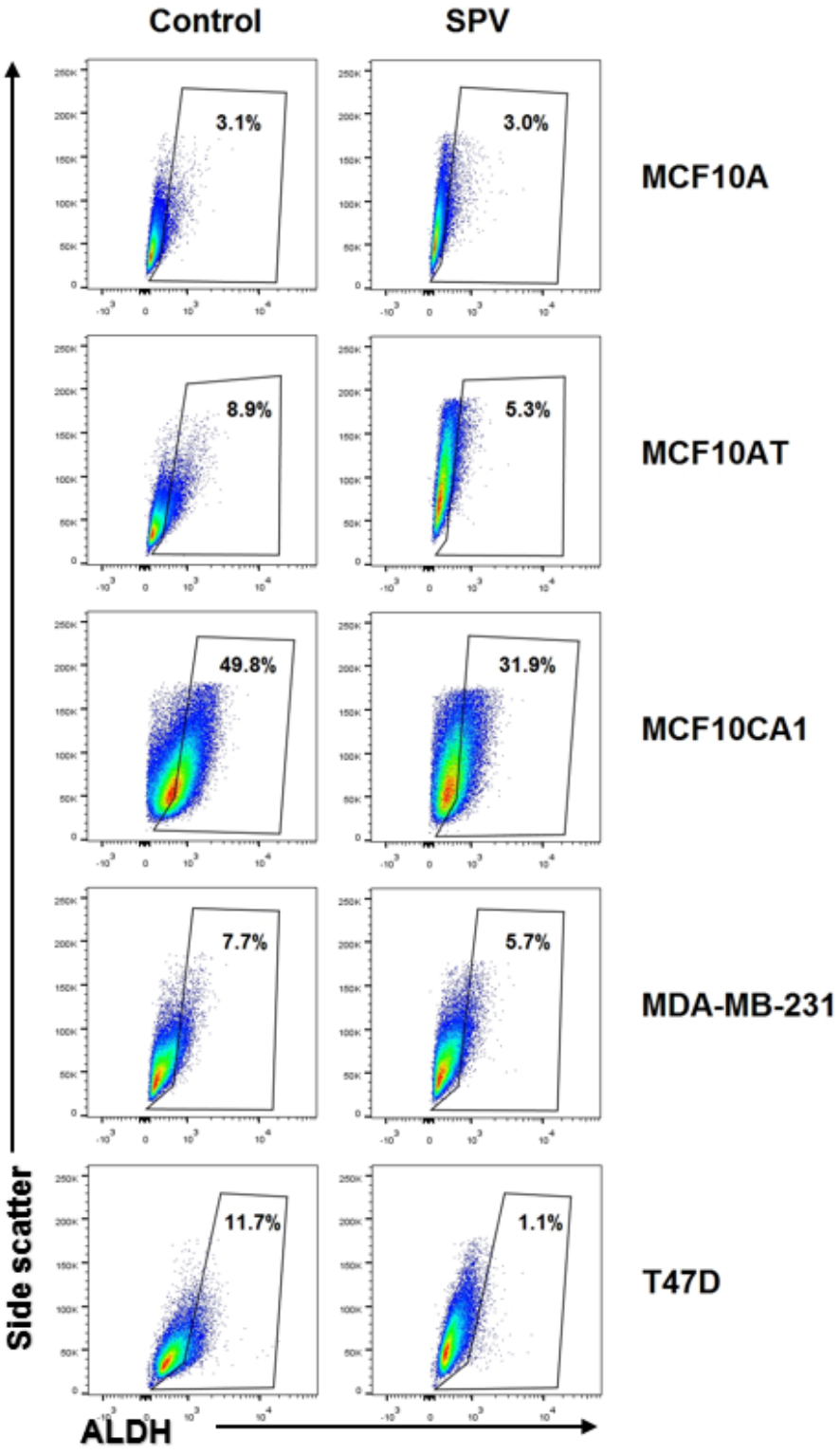


Figure S5-1 The ALDH^{high} cells were detected via ALDEFLUOR assay in untreated- and SPV-treated MCF10A, MCF10AT, MCF10CA1α, MDA-MB-231, and T47D cells

The dots inside the frame lines represent the cells with high ALDH activities. The frame lines were set based on their ALDH suppression controls.

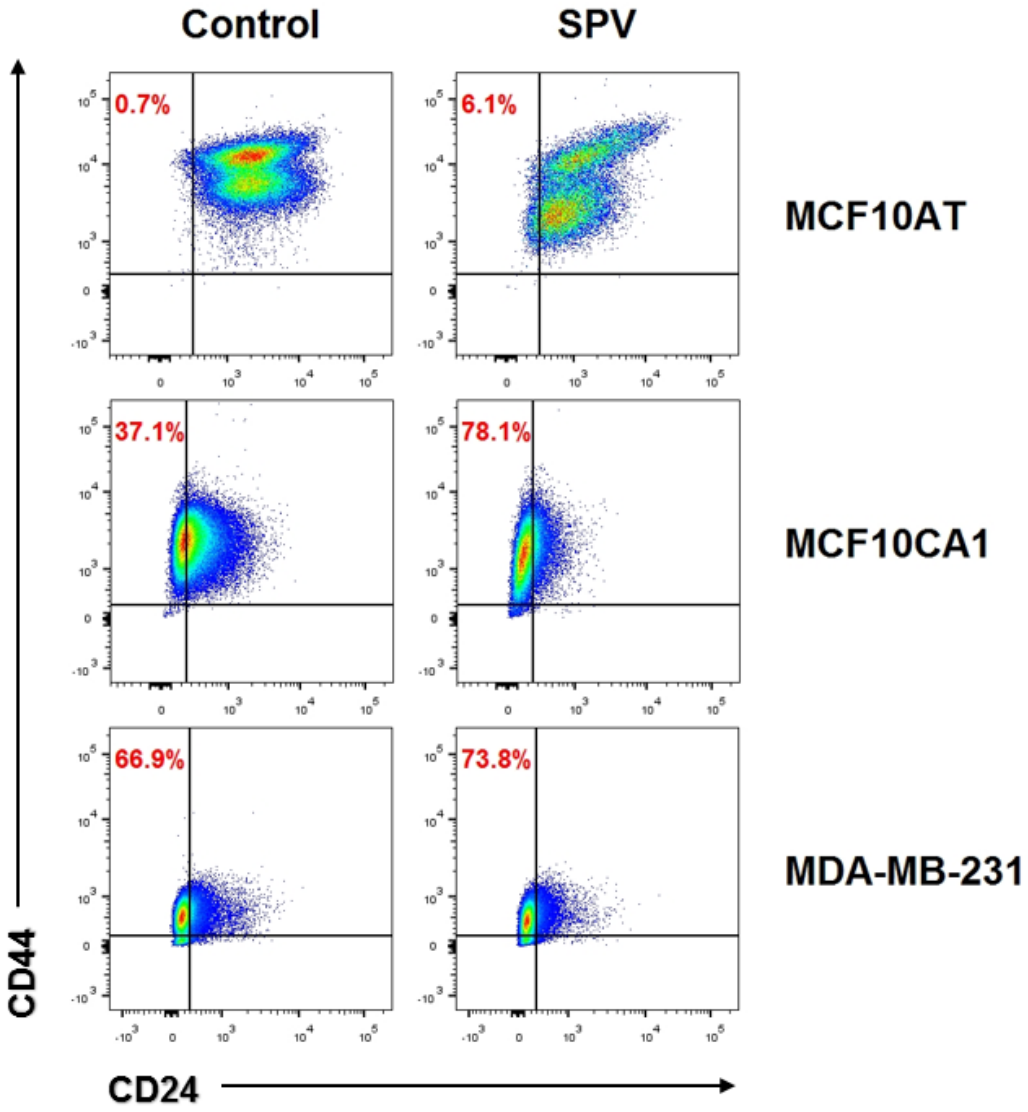


Figure S5-2 The CD44^{+}/ CD24^{-} cells were detected in untreated- and SPV-treated MCF10AT, MCF10CA1α, and MDA-MB-231 cells

The dots in the up-left corner represent the CD44^{+}/CD24^{-} *mesenchymal-like* CSCs. The frame lines were set based on their unstained controls.

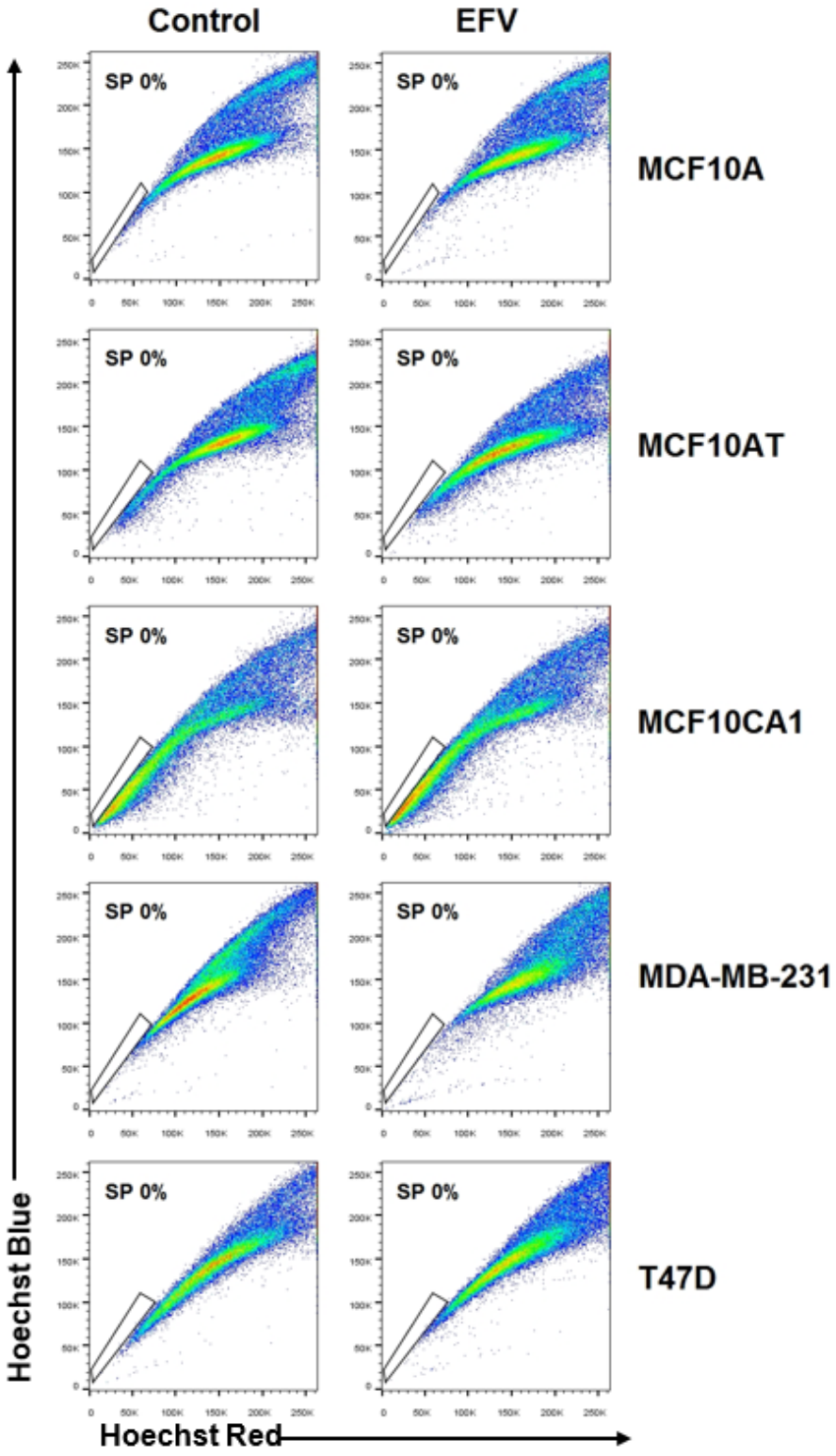


Figure S5-3 Side populations detected via the Hoechst 33342 staining method in MCF10A, MCF10AT, MCF10CA1α, MDA-MB-231, and T47D cells

The dots inside the frame lines represent the cells with high dye efflux abilities.

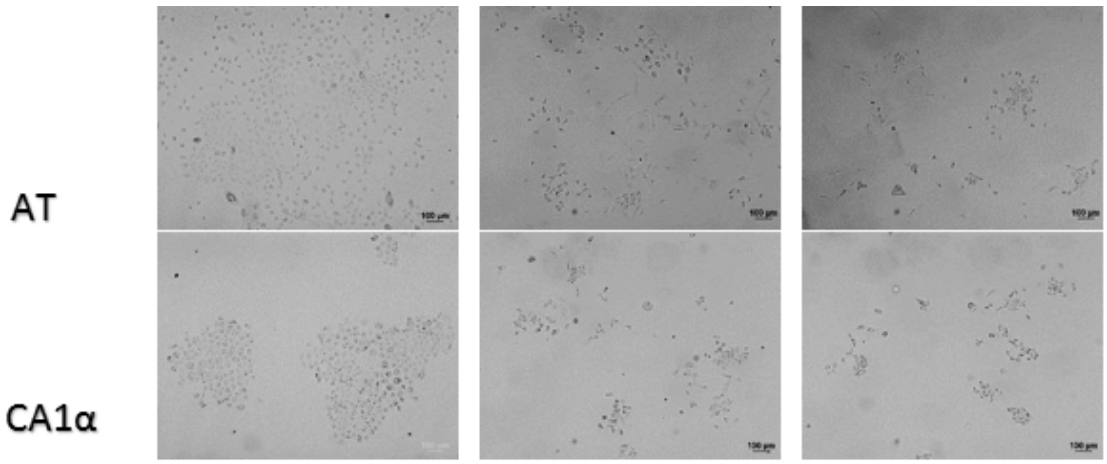


Figure S5-4 Morphological changes in MCF10AT and MCF10CA1α cells were observed after treating cells with antiretroviral drugs for 4 days

Scale bar: 100 µm

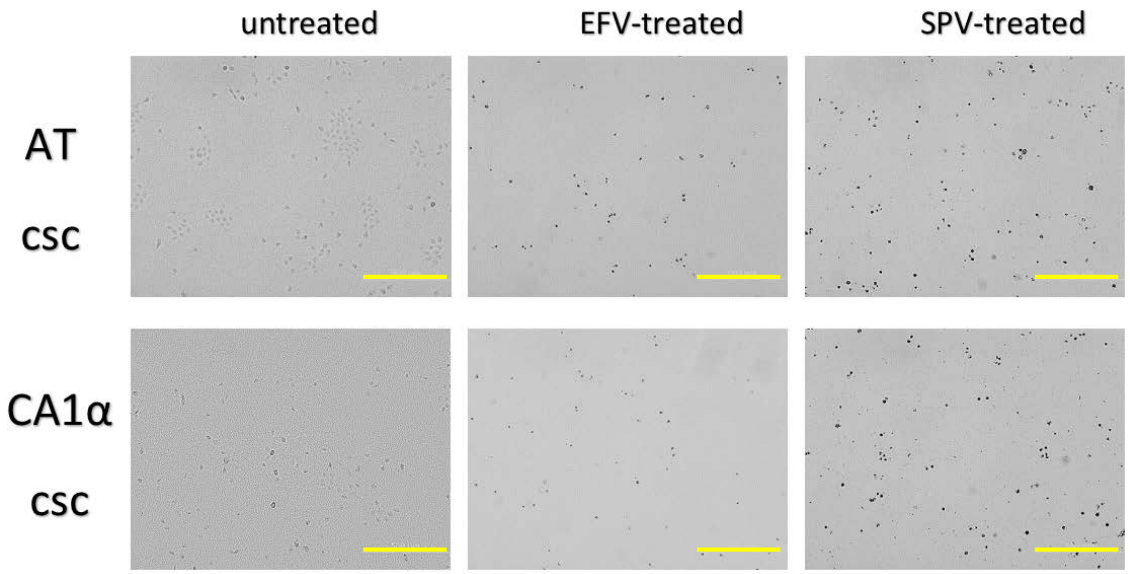


Figure S5-5 The phenotypes of sorted-CD44^{+}/CD24^{-} cells after 24 hours of drug treatment

Sorted CD44+/CD24- cells were re-cultured in the normal culture medium. When cells attached and re-grew, the cells were treated with antiretroviral drugs. After 24 hours of drug treatments, most of the drug-treated cells died. The small black dots were the dead cells (Scale bar: 500 µm)

Table S5-1 The numbers of tumorspheres being formed after 7 days of EFV treatment

Number of spheres /10^4 cell-seeded		60-100 μm			100-200 μm			200-400 μm			>400 μm			total		
MCF10A	control	9.5	5	14	14	7	16	6	2	7	1	1	1	30.5	15	38
	EFV	7.5	5	22	3.5	3	7	0.5	4	1	0	1	0	11.5	13	30
MCF10AT	control	36	47	19	27	46	46	11	11	15	9	7	3	83	111	83
	EFV	6	4	6	8	5	0	4	0	0	0	0	0	18	9	6
MCF10CA1α	control	95	79	45	91	115	53	37	33	27	3	9	8	226	236	133
	EFV	5	6	4	29	4	2	5	0	0	0	0	0	39	10	6
MDA-MB-231	control	90	93	126	179	336	315	160	92	112	63	0	9	402	428	436
	EFV	10	37	5	6	17	0	2	0	0	0	0	0	8	17	0
BT-549	control	12	27	45	38	20	50	31	24	22	34	30	32	115	101	149
	EFV	13	2	44	27	5	34	8	2	7	0	0	0	48	9	85
BT-20	control	76	129	108	105	158	174	57	50	57	25	3	4	187	211	235
	EFV	23	47	26	70	43	22	12	0	0	0	0	0	82	43	22
T47D	control	43.5	57	46	61	119	90	40	54	43	30	27	10	131	200	143
	EFV	7	3	30	7	9	16	0.5	0	0	0	0	0	7.5	9	16

Table S5-2 The Student *t*-test results for analysing the changes of antiretroviral drug-treated tumorspheres

T-Test	60-100 μm	100-200 μm	200-400 μm	>400 μm	total
MCF10A	0.28996	0.02877*	0.17280	0.091752	0.095816
MCF10AT	0.04042*	0.02546*	0.02068*	0.03478*	0.00829**
MCF10CA1α	0.02092*	0.02797*	0.00182**	0.03476*	0.01232*
MDA-MB-231	0.02287*	0.01521*	0.0126*	0.17340	0.00167**
BT-549	0.21179	0.00585**	0.00773**	0.00065**	0.00700**
BT-20	0.00870**	0.05013	0.00234**	0.13767	0.01417*
T47D	0.04191*	0.02005*	0.00455**	0.03486*	0.01705*

(one-tail, paired Student's t-test, *$p<0.05$; **$p<0.0$

Bibliography

AHMED, M. & RAHMAN, N. 2006. ATM and breast cancer susceptibility. *Oncogene,* 25**,** 5906-11.

AIHW 2017. Australian Institute of Health and Welfare 2017. *Cancer in Australia 2017,* Cancer series no.101.

ALBRECHT, W., SANTIS, M. D. & DOSSENBACH-GLANINGER, A. 2004. Testicular tumor markers: Cornerstones in the management of malignant germ cell tumors / Hoden-Tumor-marker: Eckpfeiler in der Behandlung maligner Keimzelltumoren. *LaboratoriumsMedizin.*

ALISON, M. R., GUPPY, N. J., LIM, S. M. & NICHOLSON, L. J. 2010. Finding cancer stem cells: are aldehyde dehydrogenases fit for purpose? *J Pathol,* 222**,** 335-44.

APOSTOLOVA, N., ESPLUGUES, J. V., FUNES, H. A., BLAS-GARCIA, A., ALEGRE, F. & POLO, M. 2014. Involvement of Nitric Oxide in the Mitochondrial Action of Efavirenz: A Differential Effect on Neurons and Glial Cells. *The Journal of Infectious Diseases,* 211**,** 1953-1958.

ASANGANI, I. A., RASHEED, S. A. K., NIKOLOVA, D. A., LEUPOLD, J. H., COLBURN, N. H., POST, S. & ALLGAYER, H. 2007. MicroRNA-21 (miR-21) post-transcriptionally downregulates tumor suppressor Pdcd4 and stimulates invasion, intravasation and metastasis in colorectal cancer. *Oncogene,* 27**,** 2128.

AUDET-WALSH, É., VERNIER, M., YEE, T., LAFLAMME, C., LI, S., CHEN, Y. & GIGUÈRE, V. 2018. SREBF1 Activity Is Regulated by an AR/mTOR Nuclear Axis in Prostate Cancer. *Molecular Cancer Research,* 16**,** 1396-1405.

BARDOT, B. & TOLEDO, F. 2017. Targeting MDM4 Splicing in Cancers. *Genes,* 8**,** 82.

BASAKRAN, N. S. 2015. CD44 as a potential diagnostic tumor marker. *Saudi medical journal,* 36**,** 273-279.

BASTOS, M. M., COSTA, C. C. P., BEZERRA, T. C., DA SILVA, F. C. & BOECHAT, N. 2016. Efavirenz a nonnucleoside reverse transcriptase inhibitor of first-generation: Approaches based on its medicinal chemistry. *Eur J Med Chem,* 108**,** 455-465.

BECK, C. R., COLLIER, P., MACFARLANE, C., MALIG, M., KIDD, J. M., EICHLER, E. E., BADGE, R. M. & MORAN, J. V. 2010. LINE-1 retrotransposition activity in human genomes. *Cell,* 141**,** 1159-70.

BELKAID, A., DUGUAY, S. R., OUELLETTE, R. J. & SURETTE, M. E. 2015. 17β-estradiol induces stearoyl-CoA desaturase-1 expression in estrogen receptor-positive breast cancer cells. *BMC Cancer,* 15**,** 440.

BERTHEAU, P., LEHMANN-CHE, J., VARNA, M., DUMAY, A., POIROT, B., PORCHER, R., TURPIN, E., PLASSA, L.-F., DE ROQUANCOURT, A., BOURSTYN, E., DE CREMOUX, P., JANIN, A., GIACCHETTI, S., ESPIÉ, M. & DE THÉ, H. 2013. p53 in breast cancer subtypes and new insights into response to chemotherapy. *The Breast,* 22**,** S27-S29.

BIELECKA, Z. F., MALISZEWSKA-OLEJNICZAK, K., SAFIR, I. J., SZCZYLIK, C. & CZARNECKA, A. M. 2017. Three-dimensional cell culture model utilization in cancer stem cell research. *Biol Rev Camb Philos Soc,* 92**,** 1505-1520.

BIRNIE, R., BRYCE, S. D., ROOME, C., DUSSUPT, V., DROOP, A., LANG, S. H., BERRY, P. A., HYDE, C. F., LEWIS, J. L., STOWER, M. J., MAITLAND, N. J. & COLLINS, A. T. 2008. Gene expression profiling of human prostate cancer stem cells reveals a pro-inflammatory phenotype and the importance of extracellular matrix interactions. *Genome Biology,* 9**,** R83.

BONUCCELLI, G., DE FRANCESCO, E. M., DE BOER, R., TANOWITZ, H. B. & LISANTI, M. P. 2017. NADH autofluorescence, a new metabolic biomarker for cancer stem cells: Identification of Vitamin C and CAPE as natural products targeting "stemness". *Oncotarget,* 8**,** 20667-20678.

BOYERINAS, B., PARK, S. M., HAU, A., MURMANN, A. E. & PETER, M. E. 2010. The role of let-7 in cell differentiation and cancer. *Endocr Relat Cancer,* 17**,** F19-36.

BOYLE, S. T. & KOCHETKOVA, M. 2014. Breast Cancer Stem Cells and the Immune System: Promotion, Evasion and Therapy. *Journal of Mammary Gland Biology and Neoplasia,* 19**,** 203-211.

BRAVO-CORDERO, J. J., HODGSON, L. & CONDEELIS, J. 2012. Directed cell invasion and migration during metastasis. *Current opinion in cell biology,* 24**,** 277-283.

BRITTON, K. M., EYRE, R., HARVEY, I. J., STEMKE-HALE, K., BROWELL, D., LENNARD, T. W. & MEESON, A. P. 2012. Breast cancer, side population cells and ABCG2 expression. *Cancer Lett,* 323**,** 97-105.

BRUNETTI, L., GUNDRY, M. C. & GOODELL, M. A. 2017. DNMT3A in Leukemia. *Cold Spring Harb Perspect Med,* 7.

BURNS, K. H. 2017. Transposable elements in cancer. *Nature Reviews Cancer,* 17**,** 415.

CARDENAS, M. G., OSWALD, E., YU, W., XUE, F., MACKERELL, A. D. & MELNICK, A. M. 2017. The Expanding Role of the BCL6 Oncoprotein as a Cancer Therapeutic Target. *Clinical Cancer Research,* 23**,** 885-893.

CARLINI, F., RIDOLFI, B., MOLINARI, A., PARISI, C., BOZZUTO, G., TOCCACIELI, L., FORMISANO, G., DE ORSI, D., PARADISI, S., GROBER, O. M., RAVO, M., WEISZ, A., ARCIERI, R., VELLA, S. & GAUDI, S. 2010. The reverse transcription inhibitor abacavir shows anticancer activity in prostate cancer cell lines. *PLoS One,* 5**,** e14221.

CHAKRAVARTI, B., RAVI, J. & GANJU, R. K. 2014. Cannabinoids as therapeutic agents in cancer: current status and future implications. *Oncotarget,* 5**,** 5852-5872.

CHANG, P. M., CHEN, C. H., YEH, C. C., LU, H. J., LIU, T. T., CHEN, M. H., LIU, C. Y., WU, A. T. H., YANG, M. H., TAI, S. K., MOCHLY-ROSEN, D. & HUANG, C. F. 2018. Transcriptome analysis and prognosis of ALDH isoforms in human cancer. *Sci Rep.*

CHARAFE-JAUFFRET, E., GINESTIER, C. & BIRNBAUM, D. 2009. Breast cancer stem cells: tools and models to rely on. *BMC Cancer,* 9**,** 202.

CHEAH, M. T., CHEN, J. Y., SAHOO, D., CONTRERAS-TRUJILLO, H., VOLKMER, A. K., SCHEEREN, F. A., VOLKMER, J.-P. & WEISSMAN, I. L. 2015. CD14-expressing cancer cells establish the inflammatory and proliferative tumor microenvironment in bladder cancer. *Proceedings of the National Academy of Sciences,* 112**,** 4725-4730.

CHEN, B. F. & CHAN, W. Y. 2014. The de novo DNA methyltransferase DNMT3A in development and cancer. *Epigenetics,* 9**,** 669-77.

CHEN, K., HUANG, Y.-H. & CHEN, J.-L. 2013. Understanding and targeting cancer stem cells: therapeutic implications and challenges. *Acta Pharmacologica Sinica,* 34**,** 732.

CHEN, L., DAHLSTROM, J. E., LEE, S. H. & RANGASAMY, D. 2012a. Naturally occurring endo-siRNA silences LINE-1 retrotransposons in human cells through DNA methylation. *Epigenetics,* 7**,** 758-71.

CHEN, W., DONG, J., HAIECH, J., KILHOFFER, M. C. & ZENIOU, M. 2016. Cancer Stem Cell Quiescence and Plasticity as Major Challenges in Cancer Therapy. *Stem Cells Int,* 2016**,** 1740936.

CHEN, X., LI, J., GRAY, W. H., LEHMANN, B. D., BAUER, J. A., SHYR, Y. & PIETENPOL, J. A. 2012b. TNBCtype: A Subtyping Tool for Triple-Negative Breast Cancer. *Cancer Inform,* 11**,** 147-56.

CHIANG, C.-H., HOU, M.-F. & HUNG, W.-C. 2013. Up-regulation of miR-182 by β-catenin in breast cancer increases tumorigenicity and invasiveness by targeting the matrix metalloproteinase inhibitor RECK. *Biochimica et Biophysica Acta (BBA) - General Subjects,* 1830**,** 3067-3076.

CHO, M. S., RUPAIMOOLE, R., CHOI, H.-J., NOH, K., CHEN, J., HU, Q., SOOD, A. K. & AFSHAR-KHARGHAN, V. 2016. Complement Component 3 Is Regulated by TWIST1 and Mediates Epithelial–Mesenchymal Transition. *The Journal of Immunology,* 196**,** 1412-1418.

CHOI, M., KIPPS, T. & KURZROCK, R. 2016. ATM Mutations in Cancer: Therapeutic Implications. *Molecular Cancer Therapeutics,* 15**,** 1781-1791.

CLEVERS, H. 2011. The cancer stem cell: premises, promises and challenges. *Nat Med,* 17**,** 313-9.

COATES, A., ABRAHAM, S., KAYE, S. B., SOWERBUTTS, T., FREWIN, C., FOX, R. M. & TATTERSALL, M. H. N. 1983. On the receiving end—patient perception of the side-effects of cancer chemotherapy. *European Journal of Cancer and Clinical Oncology,* 19**,** 203-208.

COLLIGNON, J., LOUSBERG, L., SCHROEDER, H. & JERUSALEM, G. 2016. Triple-negative breast cancer: treatment challenges and solutions. *Breast Cancer (Dove Med Press),* 8**,** 93-107.

CORDAUX, R. & BATZER, M. A. 2009. The impact of retrotransposons on human genome evolution. *Nat Rev Genet,* 10**,** 691-703.

CUI, Y., XING, P., WANG, Y., LIU, M., QIU, L., YING, G. & LI, B. 2017. NADPH accumulation is responsible for apoptosis in breast cancer cells induced by fatty acid synthase inhibition. *Oncotarget,* 8**,** 32576-32585.

DAI, L., HUANG, Q. & BOEKE, J. D. 2011. Effect of reverse transcriptase inhibitors on LINE-1 and Ty1 reverse transcriptase activities and on LINE-1 retrotransposition. *BMC Biochemistry,* 12**,** 18.

DAI, X., CHENG, H., BAI, Z. & LI, J. 2017. Breast Cancer Cell Line Classification and Its Relevance with Breast Tumor Subtyping. *J Cancer,* 8**,** 3131-3141.

DAVE, H., TRIVEDI, S., SHAH, M. & SHUKLA, S. 2011. Transforming growth factor beta 2: a predictive marker for breast cancer. *Indian J Exp Biol,* 49**,** 879-87.

DE CECCO, M., ITO, T., PETRASHEN, A. P., ELIAS, A. E., SKVIR, N. J., CRISCIONE, S. W., CALIGIANA, A., BROCCULI, G., ADNEY, E. M., BOEKE, J. D., LE, O., BEAUSÉJOUR, C., AMBATI, J., AMBATI, K., SIMON, M., SELUANOV, A., GORBUNOVA, V., SLAGBOOM, P. E., HELFAND, S. L., NERETTI, N. & SEDIVY, J. M. 2019. L1 drives IFN in senescent cells and promotes age-associated inflammation. *Nature,* 566**,** 73-78.

DE KONING, A. P. J., GU, W. J., CASTOE, T. A., BATZER, M. A. & POLLOCK, D. D. 2011. Repetitive Elements May Comprise Over Two-Thirds of the Human Genome. *Plos Genetics,* 7.

DENKERT, C., LIEDTKE, C., TUTT, A. & VON MINCKWITZ, G. 2017. Molecular alterations in triple-negative breast cancer—the road to new treatment strategies. *The Lancet,* 389**,** 2430-2442.

DEWAAL, D., NOGUEIRA, V., TERRY, A. R., PATRA, K. C., JEON, S.-M., GUZMAN, G., AU, J., LONG, C. P., ANTONIEWICZ, M. R. & HAY, N. 2018. Hexokinase-2 depletion inhibits glycolysis and induces oxidative phosphorylation in hepatocellular carcinoma and sensitizes to metformin. *Nature Communications,* 9**,** 446.

DEWI, D. L., ISHII, H., KANO, Y., NISHIKAWA, S., HARAGUCHI, N., SAKAI, D., SATOH, T., DOKI, Y. & MORI, M. 2011. Cancer stem cell theory in gastrointestinal malignancies: recent progress and upcoming challenges. *Journal of Gastroenterology,* 46**,** 1145.

DING, L., LAN, Z., XIONG, X., AO, H., FENG, Y., GU, H., YU, M. & CUI, Q. 2018. The Dual Role of MicroRNAs in Colorectal Cancer Progression. *International journal of molecular sciences,* 19**,** 2791.

DOHERTY, M. R., CHEON, H., JUNK, D. J., VINAYAK, S., VARADAN, V., TELLI, M. L., FORD, J. M., STARK, G. R. & JACKSON, M. W. 2017. Interferon-beta represses cancer stem cell properties in triple-negative breast cancer. *Proceedings of the National Academy of Sciences of the United States of America,* 114**,** 13792-13797.

DONG, Q., OH, J.-E., YI, J. K., KIM, R. H., SHIN, K.-H., MITSUYASU, R., PARK, N.-H. & KANG, M. K. 2013. Efavirenz induces autophagy and aberrant differentiation in normal human keratinocytes. *International journal of molecular medicine,* 31**,** 1305-1312.

DUAN, F., JIA, D., ZHAO, J., WU, W., MIN, L., SONG, S., WU, H., WANG, L., WANG, H., RUAN, Y. & GU, J. 2016. Loss of GFAT1 promotes epithelial-to-mesenchymal transition and predicts unfavorable prognosis in gastric cancer. *Oncotarget,* 7**,** 38427-38439.

EVSIKOV, A. V. & MARIN DE EVSIKOVA, C. 2016. Friend or Foe: Epigenetic Regulation of Retrotransposons in Mammalian Oogenesis and Early Development. *Yale J Biol Med,* 89**,** 487-497.

FARHAN, M., WANG, H., GAUR, U., LITTLE, P. J., XU, J. & ZHENG, W. 2017. FOXO Signaling Pathways as Therapeutic Targets in Cancer. *International journal of biological sciences,* 13**,** 815-827.

FARKASH, E. A., KAO, G. D., HORMAN, S. R. & PRAK, E. T. 2006. Gamma radiation increases endonuclease-dependent L1 retrotransposition in a cultured cell assay. *Nucleic Acids Res,* 34**,** 1196-204.

FARONATO, M., NGUYEN, V. T., PATTEN, D. K., LOMBARDO, Y., STEEL, J. H., PATEL, N., WOODLEY, L., SHOUSHA, S., PRUNERI, G., COOMBES, R. C. & MAGNANI, L. 2015. DMXL2 drives epithelial to mesenchymal transition in hormonal therapy resistant breast cancer through Notch hyper-activation. *Oncotarget,* 6**,** 22467-79.

FEDELE, M., CERCHIA, L. & CHIAPPETTA, G. 2017. The Epithelial-to-Mesenchymal Transition in Breast Cancer: Focus on Basal-Like Carcinomas. *Cancers,* 9**,** 134.

FILLMORE, C. & KUPERWASSER, C. 2007. Human breast cancer stem cell markers CD44 and CD24: enriching for cells with functional properties in mice or in man? *Breast Cancer Research,* 9.

FISHMAN, W. H. 1987. Clinical and biological significance of an isozyme tumor marker—PLAP. *Clinical Biochemistry,* 20**,** 387-392.

FLAVIN, R., PELUSO, S., NGUYEN, P. L. & LODA, M. 2010. Fatty acid synthase as a potential therapeutic target in cancer. *Future oncology (London, England),* 6**,** 551-562.

FOULKES, W. D., SMITH, I. E. & REIS-FILHO, J. S. 2010. Triple-Negative Breast Cancer. *New England Journal of Medicine,* 363**,** 1938-1948.

FREED-PASTOR, W. A. & PRIVES, C. 2012. Mutant p53: one name, many proteins. *Genes & Development,* 26**,** 1268-1286.

GAO, X., DAI, M., LI, Q., WANG, Z., LU, Y. & SONG, Z. 2017. HMGA2 regulates lung cancer proliferation and metastasis. *Thorac Cancer,* 8**,** 501-510.

GATCH, M. B., KOZLENKOV, A., HUANG, R.-Q., YANG, W., NGUYEN, J. D., GONZÁLEZ-MAESO, J., RICE, K. C., FRANCE, C. P., DILLON, G. H., FORSTER, M. J. & SCHETZ, J. A. 2013. The HIV antiretroviral drug efavirenz has LSD-like properties. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology,* 38**,** 2373-2384.

GINESTIER, C., HUR, M. H., CHARAFE-JAUFFRET, E., MONVILLE, F., DUTCHER, J., BROWN, M., JACQUEMIER, J., VIENS, P., KLEER, C. G., LIU, S., SCHOTT, A., HAYES, D., BIRNBAUM, D., WICHA, M. S. & DONTU, G. 2007. ALDH1 Is a Marker of Normal and Malignant Human Mammary Stem Cells and a Predictor of Poor Clinical Outcome. *Cell Stem Cell,* 1**,** 555-567.

GOLEBIEWSKA, A., BRONS, N. H., BJERKVIG, R. & NICLOU, S. P. 2011. Critical appraisal of the side population assay in stem cell and cancer stem cell research. *Cell Stem Cell,* 8**,** 136-47.

GOODIER, J. L., CHEUNG, L. E. & KAZAZIAN, H. H., JR. 2013. Mapping the LINE1 ORF1 protein interactome reveals associated inhibitors of human retrotransposition. *Nucleic Acids Res,* 41**,** 7401-19.

GRIFFITHS, B., LEWIS, C. A., BENSAAD, K., ROS, S., ZHANG, Q., FERBER, E. C., KONISTI, S., PECK, B., MIESS, H., EAST, P., WAKELAM, M., HARRIS, A. L. & SCHULZE, A. 2013. Sterol regulatory element binding protein-dependent regulation of lipid synthesis supports cell survival and tumor growth. *Cancer & Metabolism,* 1**,** 3.

GUINDON, J. & HOHMANN, A. G. 2011. The endocannabinoid system and cancer: therapeutic implication. *British journal of pharmacology,* 163**,** 1447-1463.

HANKER, L. C., KARN, T., HOLTRICH, U., GRAESER, M., BECKER, S., REINHARD, J., RUCKHABERLE, E., GEVENSLEBEN, H. & RODY, A. 2013. Prognostic impact of fascin-1 (FSCN1) in epithelial ovarian cancer. *Anticancer Res,* 33**,** 371-7.

HARRIS, C. R., DEWAN, A., ZUPNICK, A., NORMART, R., GABRIEL, A., PRIVES, C., LEVINE, A. J. & HOH, J. 2009. p53 responsive elements in human retrotransposons. *Oncogene,* 28**,** 3857-3865.

HARRIS, C. R., NORMART, R., YANG, Q., STEVENSON, E., HAFFTY, B. G., GANESAN, S., CORDON-CARDO, C., LEVINE, A. J. & TANG, L. H. 2010. Association of nuclear localization of a long interspersed nuclear element-1 protein in breast tumors with poor prognostic outcomes. *Genes Cancer,* 1**,** 115-24.

HAUPT, S., VIJAYAKUMARAN, R., MIRANDA, P. J., BURGESS, A., LIM, E. & HAUPT, Y. 2017. The role of MDM2 and MDM4 in breast cancer development and prevention. *Journal of molecular cell biology,* 9**,** 53-61.

HECHT, M., ERBER, S., HARRER, T., KLINKER, H., ROTH, T., PARSCH, H., FIEBIG, N., FIETKAU, R. & DISTEL, L. V. 2015. Efavirenz Has the Highest Anti-Proliferative Effect of Non-Nucleoside Reverse Transcriptase Inhibitors against Pancreatic Cancer Cells. *PLoS One,* 10**,** e0130277.

HECHT, M., HARRER, T., BÜTTNER, M., SCHWEGLER, M., ERBER, S., FIETKAU, R. & DISTEL, L. V. 2013. Cytotoxic effect of efavirenz is selective against cancer cells and associated with the cannabinoid system. *AIDS,* 27**,** 2031-2040.

HECHT, M., HARRER, T., KORBER, V., SARPONG, E. O., MOSER, F., FIEBIG, N., SCHWEGLER, M., STURZL, M., FIETKAU, R. & DISTEL, L. V. 2018. Cytotoxic effect of Efavirenz in BxPC-3 pancreatic cancer cells is based on oxidative stress and is synergistic with ionizing radiation. *Oncol Lett,* 15**,** 1728-1736.

HENRION, M. Y., PURDUE, M. P., SCELO, G., BRODERICK, P., FRAMPTON, M., RITCHIE, A., MEADE, A., LI, P., MCKAY, J., JOHANSSON, M., LATHROP, M., LARKIN, J., ROTHMAN, N., WANG, Z., CHOW, W. H., STEVENS, V. L., DIVER, W. R., ALBANES, D., VIRTAMO, J., BRENNAN, P., EISEN, T., CHANOCK, S. & HOULSTON, R. S. 2015. Common variation at 1q24.1 (ALDH9A1) is a potential risk factor for renal cancer. *PLoS One,* 10**,** e0122589.

HERAS, S. R., MACIAS, S., PLASS, M., FERNANDEZ, N., CANO, D., EYRAS, E., GARCIA-PEREZ, J. L. & CACERES, J. F. 2013. The Microprocessor controls the activity of mammalian retrotransposons. *Nat Struct Mol Biol,* 20**,** 1173-81.

HOLLIDAY, D. L. & SPEIRS, V. 2011. Choosing the right cell line for breast cancer research. *Breast Cancer Res,* 13**,** 215.

HORNSVELD, M., DANSEN, T. B., DERKSEN, P. W. & BURGERING, B. M. T. 2018. Re-evaluating the role of FOXOs in cancer. *Seminars in Cancer Biology,* 50**,** 90-100.

HOUÉDÉ, N., PIAZZA, P. V. & POURQUIER, P. 2018. LINE-1 as a therapeutic target for castration-resistant prostate cancer. *Frontiers in Bioscience-Landmark,* 23**,** 1292-1309.

HOUÉDÉ, N., PULIDO, M., MOUREY, L., JOLY, F., FERRERO, J. M., BELLERA, C., PRIOU, F., LALET, C., LAROCHE-CLARY, A., RAFFIN, M. C., ICHAS, F., PUECH, A. & PIAZZA, P. V. 2014. A Phase II Trial Evaluating the Efficacy and Safety of Efavirenz in Metastatic Castration-Resistant Prostate Cancer. *Oncologist,* 19**,** 1227-1228.

HU, B., JIANG, D., CHEN, Y., WEI, L., ZHANG, S., ZHAO, F., NI, R., LU, C. & WAN, C. 2015a. High CHMP4B expression is associated with accelerated cell proliferation and resistance to doxorubicin in hepatocellular carcinoma. *Tumour Biol,* 36**,** 2569-81.

HU, B., MA, Y., YANG, Y., ZHANG, L., HAN, H. & CHEN, J. 2018. CD44 promotes cell proliferation in non-small cell lung cancer. *Oncol Lett,* 15**,** 5627-5633.

HU, J., LEI, H., FEI, X., LIANG, S., XU, H., QIN, D., WANG, Y., WU, Y. & LI, B. 2015b. NES1/KLK10 gene represses proliferation, enhances apoptosis and down-regulates glucose metabolism of PC3 prostate cancer cells. *Scientific Reports,* 5**,** 17426.

HUDIS, C. A. & GIANNI, L. 2011. Triple-negative breast cancer: an unmet medical need. *Oncologist,* 16 Suppl 1**,** 1-11.

HWANG-VERSLUES, W. W., KUO, W.-H., CHANG, P.-H., PAN, C.-C., WANG, H.-H., TSAI, S.-T., JENG, Y.-M., SHEW, J.-Y., KUNG, J. T., CHEN, C.-H., LEE, E. Y. H. P., CHANG, K.-J. & LEE, W.-H. 2009. Multiple Lineages of Human Breast Cancer Stem/Progenitor Cells Identified by Profiling with Stem Cell Markers. *PLOS ONE,* 4**,** e8377.

IMBALZANO, K. M., TATARKOVA, I., IMBALZANO, A. N. & NICKERSON, J. A. 2009. Increasingly transformed MCF-10A cells have a progressively tumor-like phenotype in three-dimensional basement membrane culture. *Cancer Cell Int,* 9**,** 7.

ITKONEN, H. M., ENGEDAL, N., BABAIE, E., LUHR, M., GULDVIK, I. J., MINNER, S., HOHLOCH, J., TSOURLAKIS, M. C., SCHLOMM, T. & MILLS, I. G. 2014. UAP1 is overexpressed in prostate cancer and is protective against inhibitors of N-linked glycosylation. *Oncogene,* 34**,** 3744.

JAGGUPILLI, A. & ELKORD, E. 2012. Significance of CD44 and CD24 as cancer stem cell markers: an enduring ambiguity. *Clin Dev Immunol,* 2012**,** 708036.

JAMDADE, V. S., SETHI, N., MUNDHE, N. A., KUMAR, P., LAHKAR, M. & SINHA, N. 2015. Therapeutic targets of triple-negative breast cancer: a review. *Br J Pharmacol,* 172**,** 4228-37.

JIANG, Y., XIE, X., LI, Z., WANG, Z., ZHANG, Y., LING, Z., PAN, Y., WANG, Z. & CHEN, Y. 2011. Functional Cooperation of RKTG with p53 in Tumorigenesis and Epithelial–Mesenchymal Transition. *Cancer Research,* 71**,** 2959-2968.

JONES, R. B., GARRISON, K. E., WONG, J. C., DUAN, E. H., NIXON, D. F. & OSTROWSKI, M. A. 2008. Nucleoside analogue reverse transcriptase inhibitors differentially inhibit human LINE-1 retrotransposition. *PLoS One,* 3**,** e1547.

KAI, K., ARIMA, Y., KAMIYA, T. & SAYA, H. 2010. Breast cancer stem cells. *Breast Cancer,* 17**,** 80-5.

KALIMUTHO, M., PARSONS, K., MITTAL, D., LOPEZ, J. A., SRIHARI, S. & KHANNA, K. K. 2015. Targeted Therapies for Triple-Negative Breast Cancer: Combating a Stubborn Disease. *Trends Pharmacol Sci,* 36**,** 822-846.

KATT, W. P. & CERIONE, R. A. 2014. Glutaminase regulation in cancer cells: a druggable chain of events. *Drug Discovery Today,* 19**,** 450-457.

KEMP, J. R. & LONGWORTH, M. S. 2015. Crossing the LINE Toward Genomic Instability: LINE-1 Retrotransposition in Cancer. *Frontiers in chemistry,* 3**,** 68-68.

KIM, S. & ALEXANDER, C. M. 2014. Tumorsphere assay provides more accurate prediction of in vivo responses to chemotherapeutics. *Biotechnol Lett,* 36**,** 481-8.

KIM, S. J., CHOI, H., PARK, S. S., CHANG, C. & KIM, E. 2011. Stearoyl CoA desaturase (SCD) facilitates proliferation of prostate cancer cells through enhancement of androgen receptor transactivation. *Mol Cells,* 31**,** 371-7.

KOITO, A. & IKEDA, T. 2013. Intrinsic immunity against retrotransposons by APOBEC cytidine deaminases. *Front Microbiol,* 4**,** 28.

KOURI, F. M., HURLEY, L. A., DANIEL, W. L., DAY, E. S., HUA, Y., HAO, L., PENG, C.-Y., MERKEL, T. J., QUEISSER, M. A., RITNER, C., ZHANG, H., JAMES, C. D., SZNAJDER, J. I., CHIN, L., GILJOHANN, D. A., KESSLER, J. A., PETER, M. E., MIRKIN, C. A. & STEGH, A. H. 2015. miR-182 integrates apoptosis, growth, and differentiation programs in glioblastoma. *Genes & Development,* 29**,** 732-745.

KROUTTER, E. N., BELANCIO, V. P., WAGSTAFF, B. J. & ROY-ENGEL, A. M. 2009. The RNA polymerase dictates ORF1 requirement and timing of LINE and SINE retrotransposition. *PLoS Genet,* 5**,** e1000458.

KUO, C.-Y. & ANN, D. K. 2018. When fats commit crimes: fatty acid metabolism, cancer stemness and therapeutic resistance. *Cancer Communications,* 38**,** 47.

LACROIX, M., TOILLON, R.-A. & LECLERCQ, G. 2006. p53 and breast cancer, an update. 13**,** 293.

LAI, K. C., LIU, C. J., CHANG, K. W. & LEE, T. C. 2012. Depleting IFIT2 mediates atypical PKC signaling to enhance the migration and metastatic activity of oral squamous cell carcinoma cells. *Oncogene,* 32**,** 3686.

LAMPA, M., ARLT, H., HE, T., OSPINA, B., REEVES, J., ZHANG, B., MURTIE, J., DENG, G., BARBERIS, C., HOFFMANN, D., CHENG, H., POLLARD, J., WINTER, C., RICHON, V., GARCIA-ESCHEVERRIA, C., ADRIAN, F., WIEDERSCHAIN, D. & SRINIVASAN, L. 2017. Glutaminase is essential for the growth of triple-negative breast cancer cells with a deregulated glutamine metabolism pathway and its suppression synergizes with mTOR inhibition. *PLOS ONE,* 12**,** e0185092.

LECLERC, D., PHAM, D. N. T., LÉVESQUE, N., TRUONGCAO, M., FOULKES, W. D., SAPIENZA, C. & ROZEN, R. 2017. Oncogenic role of PDK4 in human colon cancer cells. *British Journal Of Cancer,* 116**,** 930.

LEE, K. K. & BOELSTERLI, U. A. 2014. Bypassing the compromised mitochondrial electron transport with methylene blue alleviates efavirenz/isoniazid-induced oxidant stress and mitochondria-mediated cell death in mouse hepatocytes. *Redox biology,* 2**,** 599-609.

LEHMANN, B. D., BAUER, J. A., CHEN, X., SANDERS, M. E., CHAKRAVARTHY, A. B., SHYR, Y. & PIETENPOL, J. A. 2011. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J Clin Invest,* 121**,** 2750-67.

LEI, R., TANG, J., ZHUANG, X., DENG, R., LI, G., YU, J., LIANG, Y., XIAO, J., WANG, H. Y., YANG, Q. & HU, G. 2014. Suppression of MIM by microRNA-182 activates RhoA and promotes breast cancer metastasis. *Oncogene,* 33**,** 1287-96.

LEVINE, A. J., TING, D. T. & GREENBAUM, B. D. 2016. P53 and the defenses against genome instability caused by transposons and repetitive elements. *BioEssays : news and reviews in molecular, cellular and developmental biology,* 38**,** 508-513.

LI, L., HUI, S., PENNELLO, G., DESTA, Z., TODD, S., NGUYEN, A. & FLOCKHART, D. 2007. Estimating a positive false discovery rate for variable selection in pharmacogenetic studies. *J Biopharm Stat,* 17**,** 883-902.

LI, L., SHAO, M., PENG, P., YANG, C., SONG, S., DUAN, F., JIA, D., ZHANG, M., ZHAO, J., ZHAO, R., WU, W., WANG, L., LI, C., WU, H., ZHANG, J., WU, X., RUAN, Y. & GU, J. 2017. High expression of GFAT1 predicts unfavorable prognosis in patients with hepatocellular carcinoma. *Oncotarget,* 8**,** 19205-19217.

LI, L., XU, N., FAN, N., MENG, Q., LUO, W., LV, L., MA, W., LIU, X., LIU, L., XU, F., WANG, H., MAO, W. & LI, Y. 2015. Upregulated KLK10 inhibits esophageal cancer proliferation and enhances cisplatin sensitivity in vitro. *Oncol Rep,* 34**,** 2325-32.

LI, Q. & LOZANO, G. 2013. Molecular pathways: targeting Mdm2 and Mdm4 in cancer therapy. *Clinical cancer research : an official journal of the American Association for Cancer Research,* 19**,** 34-41.

LI, Y., ZHANG, H., LI, Y., ZHAO, C., FAN, Y., LIU, J., LI, X., LIU, H. & CHEN, J. 2018. MiR-182 inhibits the epithelial to mesenchymal transition and metastasis of lung cancer cells by targeting the Met gene. *Molecular Carcinogenesis,* 57**,** 125-136.

LIANG, X., LAN, C., JIAO, G., FU, W., LONG, X., AN, Y., WANG, K., ZHOU, J., CHEN, T., LI, Y., XU, J., HUANG, Q., XU, B. & XIAO, J. 2017. Therapeutic inhibition of SGK1 suppresses colorectal cancer. *Experimental & molecular medicine,* 49**,** e399-e399.

LIM, L. Y., VIDNOVIC, N., ELLISEN, L. W. & LEONG, C. O. 2009. Mutant p53 mediates survival of breast cancer cells. *British Journal Of Cancer,* 101**,** 1606.

LIM, S. J., CHOI, H. G., JEON, C. K. & KIM, S. H. 2015. Increased chemoresistance to paclitaxel in the MCF10AT series of human breast epithelial cancer cells. *Oncol Rep,* 33**,** 2023-30.

LIN, J. & DING, D. 2017. The prognostic role of the cancer stem cell marker CD44 in ovarian cancer: a meta-analysis. *Cancer Cell International,* 17**,** 8.

LIN, L., HUTZEN, B., LEE, H. F., PENG, Z., WANG, W., ZHAO, C., LIN, H. J., SUN, D., LI, P. K., LI, C., KORKAYA, H., WICHA, M. S. & LIN, J. 2013. Evaluation of STAT3 signaling in ALDH+ and ALDH+/CD44+/CD24- subpopulations of breast cancer cells. *PLoS One,* 8**,** e82821.

LIU, H., LV, L. & YANG, K. 2015. Chemotherapy targeting cancer stem cells. *American journal of cancer research,* 5**,** 880-893.

LIU, S., CONG, Y., WANG, D., SUN, Y., DENG, L., LIU, Y., MARTIN-TREVINO, R., SHANG, L., MCDERMOTT, S. P., LANDIS, M. D., HONG, S., ADAMS, A., D'ANGELO, R., GINESTIER, C., CHARAFE-JAUFFRET, E., CLOUTHIER, S. G., BIRNBAUM, D., WONG, S. T., ZHAN, M., CHANG, J. C. & WICHA, M. S.

2014a. Breast cancer stem cells transition between epithelial and mesenchymal states reflective of their normal counterparts. *Stem Cell Reports,* 2**,** 78-91.

LIU, Y.-Y. 2011. Resuscitating Wild-Type p53 Expression by Disrupting Ceramide Glycosylation: A Novel Approach to Target Mutant p53 Tumors. *Cancer Research,* 71**,** 6295-6299.

LIU, Z., CHEN, X., WANG, Y., PENG, H., WANG, Y., JING, Y. & ZHANG, H. 2014b. PDK4 Protein Promotes Tumorigenesis through Activation of cAMP-response Element-binding Protein (CREB)-Ras Homolog Enriched in Brain (RHEB)-mTORC1 Signaling Cascade. *Journal of Biological Chemistry,* 289**,** 29739-29749.

LU, H. & HUANG, H. 2011. FOXO1: a potential target for human diseases. *Current drug targets,* 12**,** 1235-1244.

MASHIMA, T., SATO, S., SUGIMOTO, Y., TSURUO, T. & SEIMIYA, H. 2008. Promotion of glioma cell survival by acyl-CoA synthetase 5 under extracellular acidosis conditions. *Oncogene,* 28**,** 9.

MCDERMOTT, S. P. & WICHA, M. S. 2010. Targeting breast cancer stem cells. *Molecular Oncology,* 4**,** 404-419.

MEACHAM, C. E. & MORRISON, S. J. 2013. Tumour heterogeneity and cancer cell plasticity. *Nature,* 501**,** 328.

MELENDEZ, J., GROGG, M. & ZHENG, Y. 2011. Signaling role of Cdc42 in regulating mammalian physiology. *J Biol Chem,* 286**,** 2375-81.

MENENDEZ, J. A. & LUPU, R. 2017a. Fatty acid synthase (FASN) as a therapeutic target in breast cancer. *Expert Opin Ther Targets,* 21**,** 1001-1016.

MENENDEZ, J. A. & LUPU, R. 2017b. Fatty acid synthase regulates estrogen receptor-α signaling in breast cancer cells. *Oncogenesis,* 6**,** e299.

MIGITA, T., TAKAYAMA, K. I., URANO, T., OBINATA, D., IKEDA, K., SOGA, T., TAKAHASHI, S. & INOUE, S. 2017. ACSL3 promotes intratumoral steroidogenesis in prostate cancer cells. *Cancer Sci,* 108**,** 2011-2021.

MISHRA, A. K., PARISH, C. R., WONG, M. L., LICINIO, J. & BLACKBURN, A. C. 2017. Leptin signals via TGFB1 to promote metastatic potential and stemness in breast cancer. *PLoS One,* 12**,** e0178454.

MOGHBELI, M., MOGHBELI, F., FORGHANIFARD, M. M. & ABBASZADEGAN, M. R. 2014. Cancer stem cell detection and isolation. *Med Oncol,* 31**,** 69.

MORISHITA, A., ZAIDI, M. R., MITORO, A., SANKARASHARMA, D., SZABOLCS, M., OKADA, Y., D'ARMIENTO, J. & CHADA, K. 2013. HMGA2 Is a Driver of Tumor Metastasis. *Cancer Research,* 73**,** 4289-4299.

MOSERLE, L., GHISI, M., AMADORI, A. & INDRACCOLO, S. 2010. Side population and cancer stem cells: therapeutic implications. *Cancer Lett,* 288**,** 1-9.

MULLER, P. A. J. & VOUSDEN, K. H. 2014. Mutant p53 in cancer: new functions and therapeutic opportunities. *Cancer cell,* 25**,** 304-317.

NAKANISHI, T., CHUMSRI, S., KHAKPOUR, N., BRODIE, A. H., LEYLAND-JONES, B., HAMBURGER, A. W., ROSS, D. D. & BURGER, A. M. 2010. Side-population cells in luminal-type breast cancer have tumour-initiating cell properties, and are regulated by HER2 expression and signalling. *Br J Cancer,* 102**,** 815-26.

NAKARAI, C., OSAWA, K., MATSUBARA, N., IKEUCHI, H., YAMANO, T., OKAMURA, S., KAMOSHIDA, S., TSUTOU, A., TAKAHASHI, J., EJIRI, K., HIROTA, S., TOMITA, N. & KIDO, Y. 2012. Significance of ELF3 mRNA expression for detection of lymph node metastases of colorectal cancer. *Anticancer Res,* 32**,** 3753-8.

OHMS, S. & RANGASAMY, D. 2014. Silencing of LINE-1 retrotransposons contributes to variation in small noncoding RNA expression in human cancer cells. *Oncotarget,* 5**,** 4103-4117.

OLMEDILLAS LÓPEZ, S., GARCIA-ARRANZ, M., GARCIA-OLMO, D. & LIRAS, A. 2016. Preliminary study on non-viral transfection of F9 (factor IX) gene by nucleofection in human adipose-derived mesenchymal stem cells. *PeerJ,* 4**,** e1907.

ONITILO, A. A., ENGEL, J. M., GREENLEE, R. T. & MUKESH, B. N. 2009. Breast cancer subtypes based on ER/PR and Her2 expression: comparison of clinicopathologic features and survival. *Clin Med Res,* 7**,** 4-13.

OU, B., ZHAO, J., GUAN, S., WANGPU, X., ZHU, C., ZONG, Y., MA, J., SUN, J., ZHENG, M., FENG, H. & LU, A. 2016. Plk2 promotes tumor growth and inhibits apoptosis by targeting Fbxw7/Cyclin E in colorectal cancer. *Cancer Letters,* 380**,** 457-466.

PADANAD, M. S., KONSTANTINIDOU, G., VENKATESWARAN, N., MELEGARI, M., RINDHE, S., MITSCHE, M., YANG, C., BATTEN, K., HUFFMAN, K. E., LIU, J., TANG, X., RODRIGUEZ-CANALES, J., KALHOR, N., SHAY, J. W., MINNA, J. D., MCDONALD, J., WISTUBA, II, DEBERARDINIS, R. J. & SCAGLIONI, P. P. 2016. Fatty Acid Oxidation Mediated by Acyl-CoA Synthetase Long Chain 3 Is Required for Mutant KRAS Lung Tumorigenesis. *Cell Rep,* 16**,** 1614-1628.

PADDISON, P. J., CLEARY, M., SILVA, J. M., CHANG, K., SHETH, N., SACHIDANANDAM, R. & HANNON, G. J. 2004. Cloning of short hairpin RNAs for gene knockdown in mammalian cells. *Nature Methods,* 1**,** 163.

PARRALES, A. & IWAKUMA, T. 2015. Targeting Oncogenic Mutant p53 for Cancer Therapy. *Frontiers in Oncology,* 5.

PATNALA, R., LEE, S. H., DAHLSTROM, J. E., OHMS, S., CHEN, L., DHEEN, S. T. & RANGASAMY, D. 2014. Inhibition of LINE-1 retrotransposon-encoded reverse transcriptase modulates the expression of cell differentiation genes in breast cancer cells. *Breast Cancer Res Treat,* 143**,** 239-53.

PATRA, K. C., WANG, Q., BHASKAR, P. T., MILLER, L., WANG, Z., WHEATON, W., CHANDEL, N., LAAKSO, M., MULLER, W. J., ALLEN, E. L., JHA, A. K., SMOLEN, G. A., CLASQUIN, M. F., ROBEY, B. & HAY, N. 2013. Hexokinase 2 is required for tumor initiation and maintenance and its systemic deletion is therapeutic in mouse models of cancer. *Cancer cell,* 24**,** 213-228.

PEĆINA-SLAUS, N. 2003. Tumor suppressor gene E-cadherin and its role in normal and malignant cells. *Cancer cell international,* 3**,** 17-17.

PECK, B., SCHUG, Z. T., ZHANG, Q., DANKWORTH, B., JONES, D. T., SMETHURST, E., PATEL, R., MASON, S., JIANG, M., SAUNDERS, R., HOWELL, M., MITTER, R., SPENCER-DENE, B., STAMP, G., MCGARRY, L., JAMES, D., SHANKS, E., ABOAGYE, E. O., CRITCHLOW, S. E., LEUNG, H. Y., HARRIS, A. L., WAKELAM, M. J. O., GOTTLIEB, E. & SCHULZE, A. 2016. Inhibition of fatty acid desaturation is detrimental to cancer cell survival in metabolically compromised environments. *Cancer Metab,* 4**,** 6.

PEDDIGARI, S., LI, P. W., RABE, J. L. & MARTIN, S. L. 2013. hnRNPL and nucleolin bind LINE-1 RNA and function as host factors to modulate retrotransposition. *Nucleic Acids Res,* 41**,** 575-85.

PENG, Y. P., ZHU, Y., YIN, L. D., WEI, J. S., LIU, X. C., ZHU, X. L. & MIAO, Y. 2018. PIK3R3 Promotes Metastasis of Pancreatic Cancer via ZEB1 Induced Epithelial-Mesenchymal Transition. *Cellular Physiology and Biochemistry,* 46**,** 1930-1938.

PERERA, F., TANG, W. Y., HERBSTMAN, J., TANG, D., LEVIN, L., MILLER, R. & HO, S. M. 2009. Relation of DNA methylation of 5'-CpG island of ACSL3 to transplacental exposure to airborne polycyclic aromatic hydrocarbons and childhood asthma. *PLoS One,* 4**,** e4488.

PERONI, R. N., DI GENNARO, S. S., HOCHT, C., CHIAPPETTA, D. A., RUBIO, M. C., SOSNIK, A. & BRAMUGLIA, G. F. 2011. Efavirenz is a substrate and in turn modulates the expression of the efflux transporter ABCG2/BCRP in the gastrointestinal tract of the rat. *Biochemical Pharmacology,* 82**,** 1227-1233.

PFEFFER, S. R., YANG, C. H. & PFEFFER, L. M. 2015. The Role of miR-21 in Cancer. *Drug Development Research,* 76**,** 270-277.

PURNELL, P. R. & FOX, H. S. 2014. Efavirenz induces neuronal autophagy and mitochondrial alterations. *The Journal of pharmacology and experimental therapeutics,* 351**,** 250-258.

RANIGA, K. & LIANG, C. 2018. Interferons: Reprogramming the Metabolic Network against Viral Infection. *Viruses,* 10**,** 36.

RAO, S. R., SNAITH, A. E., MARINO, D., CHENG, X., LWIN, S. T., ORRISS, I. R., HAMDY, F. C. & EDWARDS, C. M. 2016. Tumour-derived alkaline phosphatase regulates tumour growth, epithelial plasticity and disease-free survival in metastatic prostate cancer. *British Journal Of Cancer,* 116**,** 227.

RAVICHANDRAN, K. S. 2001. Signaling via Shc family adapter proteins. *Oncogene,* 20**,** 6322.

REDDY, K. B. 2011. Triple-negative breast cancers: an updated review on treatment options. *Curr Oncol,* 18**,** e173-9.

RODIĆ, N. & BURNS, K. H. 2013. Long interspersed element-1 (LINE-1): passenger or driver in human neoplasms? *PLoS Genet,* 9**,** e1003402.

RODIĆ, N., SHARMA, R., SHARMA, R., ZAMPELLA, J., DAI, L., TAYLOR, M. S., HRUBAN, R. H., IACOBUZIO-DONAHUE, C. A., MAITRA, A., TORBENSON, M. S., GOGGINS, M., SHIH IE, M., DUFFIELD, A. S., MONTGOMERY, E. A., GABRIELSON, E., NETTO, G. J., LOTAN, T. L., DE MARZO, A. M., WESTRA, W., BINDER, Z. A., ORR, B. A., GALLIA, G. L., EBERHART, C. G., BOEKE, J. D., HARRIS, C. R. & BURNS, K. H. 2014. Long interspersed element-1 protein expression is a hallmark of many human cancers. *Am J Pathol,* 184**,** 1280-6.

RODRIGUEZ SALAS, N., GONZALEZ GONZALEZ, E. & GAMALLO AMAT, C. 2010. Breast cancer stem cell hypothesis: clinical relevance (answering breast cancer clinical features). *Clin Transl Oncol,* 12**,** 395-400.

SACHIDANANDAM, R., WEISSMAN, D., SCHMIDT, S. C., KAKOL, J. M., STEIN, L. D., MARTH, G., SHERRY, S., MULLIKIN, J. C., MORTIMORE, B. J., WILLEY, D. L., HUNT, S. E., COLE, C. G., COGGILL, P. C., RICE, C. M., NING, Z., ROGERS, J., BENTLEY, D. R., KWOK, P. Y., MARDIS, E. R., YEH, R. T., SCHULTZ, B., COOK, L., DAVENPORT, R., DANTE, M., FULTON, L., HILLIER, L., WATERSTON, R. H., MCPHERSON, J. D., GILMAN, B., SCHAFFNER, S., VAN ETTEN, W. J., REICH, D., HIGGINS, J., DALY, M. J., BLUMENSTIEL, B., BALDWIN, J., STANGE-THOMANN, N., ZODY, M. C., LINTON, L., LANDER, E. S., ALTSHULER, D. & INTERNATIONAL, S. N. P. M. W. G. 2001. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature,* 409**,** 928-33.

SAHAB, Z. J., HALL, M. D., ME SUNG, Y., DAKSHANAMURTHY, S., JI, Y., KUMAR, D. & BYERS, S. W. 2011. Tumor suppressor RARRES1 interacts with cytoplasmic carboxypeptidase AGBL2 to regulate the α-tubulin tyrosination cycle. *Cancer research,* 71**,** 1219-1228.

SATELLI, A. & LI, S. 2011. Vimentin in cancer and its potential as a molecular target for cancer therapy. *Cellular and Molecular Life Sciences,* 68**,** 3033-3046.

SBARDELLA, G., MAI, A., BARTOLINI, S., CASTELLANO, S., CIRILLI, R., ROTILI, D., MILITE, C., SANTORIELLO, M., ORLANDO, S., SCIAMANNA, I., SERAFINO, A., LAVIA, P. & SPADAFORA, C. 2011. Modulation of cell differentiation, proliferation, and tumor growth by dihydrobenzyloxopyrimidine non-nucleoside reverse transcriptase inhibitors. *J Med Chem,* 54**,** 5927-36.

SCHIEBER, M. & CHANDEL, NAVDEEP S. 2014. ROS Function in Redox Signaling and Oxidative Stress. *Current Biology,* 24**,** R453-R462.

SCHMID, P., ADAMS, S., RUGO, H. S., SCHNEEWEISS, A., BARRIOS, C. H., IWATA, H., DIÉRAS, V., HEGG, R., IM, S.-A., SHAW WRIGHT, G., HENSCHEL, V., MOLINERO, L., CHUI, S. Y., FUNKE, R., HUSAIN, A., WINER, E. P., LOI, S. & EMENS, L. A. 2018. Atezolizumab and Nab-Paclitaxel in Advanced Triple-Negative Breast Cancer. *New England Journal of Medicine,* 379**,** 2108-2121.

SCIAMANNA, I., DE LUCA, C. & SPADAFORA, C. 2016. The Reverse Transcriptase Encoded by LINE-1 Retrotransposons in the Genesis, Progression, and Therapy of Cancer. *Front Chem,* 4**,** 6.

SCIAMANNA, I., GUALTIERI, A., PIAZZA, P. F. & SPADAFORA, C. 2014. Regulatory roles of LINE-1-encoded reverse transcriptase in cancer onset and progression. *Oncotarget,* 5**,** 8039-51.

SCIAMANNA, I., LANDRISCINA, M., PITTOGGI, C., QUIRINO, M., MEARELLI, C., BERALDI, R., MATTEI, E., SERAFINO, A., CASSANO, A., SINIBALDI-VALLEBONA, P., GARACI, E., BARONE, C. & SPADAFORA, C. 2005. Inhibition of endogenous reverse transcriptase antagonizes human tumor growth. *Oncogene,* 24**,** 3923-31.

SCIAMANNA, I., SINIBALDI-VALLEBONA, P., SERAFINO, A. & SPADAFORA, C. 2018. LINE-1-encoded reverse Transcriptase as a target in cancer therapy. *Front Biosci (Landmark Ed),* 23**,** 1360-1369.

SEIFARTH, W., FRANK, O., ZEILFELDER, U., SPIESS, B., GREENWOOD, A. D., HEHLMANN, R. & LEIB-MÖSCH, C. 2005. Comprehensive Analysis of Human Endogenous Retrovirus Transcriptional Activity in Human Tissues with a Retrovirus-Specific Microarray. *Journal of Virology,* 79**,** 341-352.

SHARMA, L. K., FANG, H., LIU, J., VARTAK, R., DENG, J. & BAI, Y. 2011. Mitochondrial respiratory complex I dysfunction promotes tumorigenesis through ROS alteration and AKT activation. *Human Molecular Genetics,* 20**,** 4605-4616.

SHERIDAN, C., KISHIMOTO, H., FUCHS, R. K., MEHROTRA, S., BHAT-NAKSHATRI, P., TURNER, C. H., GOULET, R., JR., BADVE, S. & NAKSHATRI, H. 2006. CD44+/CD24- breast cancer cells exhibit enhanced invasive properties: an early step necessary for metastasis. *Breast Cancer Res,* 8**,** R59.

SHI, B., SHARIFI, H. J., DIGRIGOLI, S., KINNETZ, M., MELLON, K., HU, W. & DE NORONHA, C. M. C. 2018. Inhibition of HIV early replication by the p53 and its downstream gene p21. *Virology Journal,* 15**,** 53.

SIMARD, E. P. & ENGELS, E. A. 2010. Cancer as a Cause of Death among People with AIDS in the United States. *Clinical Infectious Diseases,* 51**,** 957-962.

SINGH, A. & SETTLEMAN, J. 2010. EMT, cancer stem cells and drug resistance: an emerging axis of evil in the war on cancer. *Oncogene,* 29**,** 4741-51.

SLEIRE, L., FORDE, H. E., NETLAND, I. A., LEISS, L., SKEIE, B. S. & ENGER, P. O. 2017. Drug repurposing in cancer. *Pharmacol Res,* 124**,** 74-91.

SNYDER, V., REED-NEWMAN, T. C., ARNOLD, L., THOMAS, S. M. & ANANT, S. 2018. Cancer Stem Cell Metabolism and Potential Therapeutic Targets. *Front Oncol,* 8**,** 203.

SOUSSI, T. & WIMAN, K. G. 2015. TP53: an oncogene in disguise. *Cell Death And Differentiation,* 22**,** 1239.

SPADAFORA, C. 2004. Endogenous reverse transcriptase: a mediator of cell proliferation and differentiation. *Cytogenetic and Genome Research,* 105**,** 346-350.

SPINELLI, J. B. & HAIGIS, M. C. 2018. The multifaceted contributions of mitochondria to cellular metabolism. *Nature Cell Biology,* 20**,** 745-754.

STOREY, J. D. 2003. The positive false discovery rate: a Bayesian interpretation and the q -value. *Ann. Statist.,* 31**,** 2013-2035.

SUN, Y., HE, W., LUO, M., ZHOU, Y., CHANG, G., REN, W., WU, K., LI, X., SHEN, J., ZHAO, X. & HU, Y. 2015. SREBP1 regulates tumorigenesis and prognosis of pancreatic cancer through targeting lipid metabolism. *Tumor Biology,* 36**,** 4133-4141.

SZKLARCZYK, D., MORRIS, J. H., COOK, H., KUHN, M., WYDER, S., SIMONOVIC, M., SANTOS, A., DONCHEVA, N. T., ROTH, A., BORK, P., JENSEN, L. J. & VON MERING, C. 2017. The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. *Nucleic Acids Research,* 45**,** D362-D368.

TALARICO, C., DATTILO, V., D'ANTONA, L., MENNITI, M., BIANCO, C., ORTUSO, F., ALCARO, S., SCHENONE, S., PERROTTI, N. & AMATO, R. 2016. SGK1, the New Player in the Game of Resistance: Chemo-Radio Molecular Target and Strategy for Inhibition. *Cellular Physiology and Biochemistry,* 39**,** 1863-1876.

TAMHANE, A. C. & DUNLOP, D. D. 2000. *Statistics and data analysis : from elementary to intermediate,* Upper Saddle River, NJ, Prentice Hall.

TAN, M. J., TEO, Z., SNG, M. K., ZHU, P. & TAN, N. S. 2012. Emerging Roles of Angiopoietin-like 4 in Human Cancer. *Molecular Cancer Research,* 10**,** 677-688.

TANG, D. G. 2012. Understanding cancer stem cell heterogeneity and plasticity. *Cell Research,* 22**,** 457.

TIRINATO, L., PAGLIARI, F., LIMONGI, T., MARINI, M., FALQUI, A., SECO, J., CANDELORO, P., LIBERALE, C. & DI FABRIZIO, E. 2017. An Overview of Lipid Droplets in Cancer and Cancer Stem Cells. *Stem Cells Int,* 2017**,** 1656053.

TURNER, N., MORETTI, E., SICLARI, O., MIGLIACCIO, I., SANTARPIA, L., D'INCALCI, M., PICCOLO, S., VERONESI, A., ZAMBELLI, A., DEL SAL, G. & DI LEO, A. 2013. Targeting triple negative breast cancer: Is p53 the answer? *Cancer Treatment Reviews,* 39**,** 541-550.

VARNA, M., BOUSQUET, G., PLASSA, L. F., BERTHEAU, P. & JANIN, A. 2011. TP53 status and response to treatment in breast cancers. *J Biomed Biotechnol,* 2011**,** 284584.

VELASCO-VELAZQUEZ, M. A., HOMSI, N., DE LA FUENTE, M. & PESTELL, R. G. 2012. Breast cancer stem cells. *Int J Biochem Cell Biol,* 44**,** 573-7.

VIRANI, S., COLACINO, J. A., KIM, J. H. & ROZEK, L. S. 2012. Cancer epigenetics: a brief review. *ILAR J,* 53**,** 359-69.

WAGSTAFF, B. J., BARNERSSOI, M. & ROY-ENGEL, A. M. 2011. Evolutionary conservation of the functional modularity of primate and murine LINE-1 elements. *PLoS One,* 6**,** e19672.

WANG, C.-Q., TANG, C.-H., WANG, Y., JIN, L., WANG, Q., LI, X., HU, G.-N., HUANG, B.-F., ZHAO, Y.-M. & SU, C.-M. 2017a. FSCN1 gene polymorphisms: biomarkers for the development and progression of breast cancer. *Scientific Reports,* 7**,** 15887.

WANG, D., HUANG, J. & HU, Z. 2012. RNA helicase DDX5 regulates microRNA expression and contributes to cytoskeletal reorganization in basal breast cancer cells. *Mol Cell Proteomics,* 11**,** M111 011932.

WANG, G., YANG, X., LI, C., CAO, X., LUO, X. & HU, J. 2014a. PIK3R3 Induces Epithelial-to-Mesenchymal Transition and Promotes Metastasis in Colorectal Cancer. *Molecular Cancer Therapeutics,* 13**,** 1837-1847.

WANG, H., WANG, L., ZHANG, Y., WANG, J., DENG, Y. & LIN, D. 2016. Inhibition of glycolytic enzyme hexokinase II (HK2) suppresses lung tumor growth. *Cancer Cell International,* 16**,** 9.

WANG, H., YU, Z., HUO, S., CHEN, Z., OU, Z., MAI, J., DING, S. & ZHANG, J. 2018. Overexpression of ELF3 facilitates cell growth and metastasis through PI3K/Akt and ERK signaling pathways in non-small cell lung cancer. *The International Journal of Biochemistry & Cell Biology,* 94**,** 98-106.

WANG, J. L., CHEN, Z. F., CHEN, H. M., WANG, M. Y., KONG, X., WANG, Y. C., SUN, T. T., HONG, J., ZOU, W., XU, J. & FANG, J. Y. 2014b. Elf3 drives β-catenin transactivation and associates with poor prognosis in colorectal cancer. *Cell Death &Amp; Disease,* 5**,** e1263.

WANG, J. P. & HIELSCHER, A. 2017. Fibronectin: How Its Aberrant Expression in Tumors May Improve Therapeutic Targeting. *Journal of Cancer,* 8**,** 674-682.

WANG, L., ZHANG, T., WANG, L., CAI, Y., ZHONG, X., HE, X., HU, L., TIAN, S., WU, M., HUI, L., ZHANG, H. & GAO, P. 2017b. Fatty acid synthesis is critical for stem cell pluripotency via promoting mitochondrial fission. *The EMBO Journal,* 36**,** 1330-1347.

WANG, Q., DU, X., ZHOU, B., LI, J., LU, W., CHEN, Q. & GAO, J. 2017c. Mitochondrial dysfunction is responsible for fatty acid synthase inhibition-induced apoptosis in breast cancer cells by PdpaMn. *Biomedicine & Pharmacotherapy,* 96**,** 396-403.

WANG, Z., JIANG, Y., GUAN, D., LI, J., YIN, H., PAN, Y., XIE, D. & CHEN, Y. 2013. Critical Roles of p53 in Epithelial-Mesenchymal Transition and Metastasis of Hepatocellular Carcinoma Cells. *PLOS ONE,* 8**,** e72846.

WEISS, J., ROSE, J., STORCH, C. H., KETABI-KIYANVASH, N., SAUER, A., HAEFELI, W. E. & EFFERTH, T. 2007. Modulation of human BCRP (ABCG2) activity by anti-HIV drugs. *Journal of Antimicrobial Chemotherapy,* 59**,** 238-245.

WEN, Y.-A., XIONG, X., ZAYTSEVA, Y. Y., NAPIER, D. L., VALLEE, E., LI, A. T., WANG, C., WEISS, H. L., EVERS, B. M. & GAO, T. 2018. Downregulation of SREBP inhibits tumor growth and initiation by altering cellular metabolism in colon cancer. *Cell Death & Disease,* 9**,** 265.

WHEATON, W. W., WEINBERG, S. E., HAMANAKA, R. B., SOBERANES, S., SULLIVAN, L. B., ANSO, E., GLASAUER, A., DUFOUR, E., MUTLU, G. M., BUDIGNER, G. R. S. & CHANDEL, N. S. 2014. Metformin inhibits mitochondrial complex I of cancer cells to reduce tumorigenesis. *eLife,* 3**,** e02242.

WOOLBRIGHT, B. L., CHOUDHARY, D., MIKHALYUK, A., TRAMMEL, C., SHANMUGAM, S., ABBOTT, E., PILBEAM, C. C. & TAYLOR, J. A. 2018. The Role of Pyruvate Dehydrogenase Kinase-4 (PDK4) in Bladder Cancer and Chemoresistance. *Molecular Cancer Therapeutics,* 17**,** 2004-2012.

WU, C. & ALMAN, B. A. 2008. Side population cells in human cancers. *Cancer Lett,* 268**,** 1-9.

WU, D., SANIN, DAVID E., EVERTS, B., CHEN, Q., QIU, J., BUCK, MICHAEL D., PATTERSON, A., SMITH, AMBER M., CHANG, C.-H., LIU, Z., ARTYOMOV, MAXIM N., PEARCE, ERIKA L., CELLA, M. & PEARCE, EDWARD J. 2016a. Type 1 Interferons Induce Changes in Core Metabolism that Are Critical for Immune Function. *Immunity,* 44**,** 1325-1336.

WU, J., HU, L., WU, F., ZOU, L. & HE, T. 2017. Poor prognosis of hexokinase 2 overexpression in solid tumors of digestive system: a meta-analysis. *Oncotarget,* 8**,** 32332-32344.

WU, J., ZHANG, S., SHAN, J., HU, Z., LIU, X., CHEN, L., REN, X., YAO, L., SHENG, H., LI, L., ANN, D., YEN, Y., WANG, J. & WANG, X. 2016b. Elevated HMGA2 expression is associated with cancer aggressiveness and predicts poor outcome in breast cancer. *Cancer Letters,* 376**,** 284-292.

WU, Q., LIU, X., YAN, H., HE, Y.-H., YE, S., CHENG, X.-W., ZHU, G.-L., WU, W.-Y., WANG, X.-N., KONG, X.-J., XU, X.-C., LOBIE, P. E., ZHU, T. & WU, Z.-S. 2014. B-cell lymphoma 6 protein stimulates oncogenicity of human breast cancer cells. *BMC Cancer,* 14**,** 418.

XUE, B. & HE, L. 2014. An expanding universe of the non-coding genome in cancer biology. *Carcinogenesis,* 35**,** 1209-16.

YAN, L.-X., HUANG, X.-F., SHAO, Q., HUANG, M.-Y., DENG, L., WU, Q.-L., ZENG, Y.-X. & SHAO, J.-Y. 2008. MicroRNA miR-21 overexpression in human breast cancer is associated with advanced clinical stage, lymph node metastasis and patient poor prognosis. *RNA (New York, N.Y.),* 14**,** 2348-2360.

YANG, C., PENG, P., LI, L., SHAO, M., ZHAO, J., WANG, L., DUAN, F., SONG, S., WU, H., ZHANG, J., ZHAO, R., JIA, D., ZHANG, M., WU, W., LI, C., RONG, Y., ZHANG, L., RUAN, Y. & GU, J. 2016. High expression of GFAT1 predicts poor prognosis in patients with pancreatic cancer. *Scientific Reports,* 6**,** 39044.

YI, M., LI, J., CHEN, S., CAI, J., BAN, Y., PENG, Q., ZHOU, Y., ZENG, Z., PENG, S., LI, X., XIONG, W., LI, G. & XIANG, B. 2018. Emerging role of lipid metabolism alterations in Cancer stem cells. *J Exp Clin Cancer Res,* 37**,** 118.

YU, C. C., FURUKAWA, M., KOBAYASHI, K., SHIKISHIMA, C., CHA, P. C., SESE, J., SUGAWARA, H., IWAMOTO, K., KATO, T., ANDO, J. & TODA, T. 2012. Genome-wide DNA methylation and gene expression analyses of monozygotic twins discordant for intelligence levels. *PLoS One,* 7**,** e47081.

YUE, D. & QIN, X. 2019. miR-182 regulates trastuzumab resistance by targeting MET in breast cancer cells. *Cancer Gene Therapy,* 26**,** 1-10.

ZEICHNER, S. B., TERAWAKI, H. & GOGINENI, K. 2016. A Review of Systemic Treatment in Metastatic Triple-Negative Breast Cancer. *Breast cancer : basic and clinical research,* 10**,** 25-36.

ZHANG, H., CAI, K., WANG, J., WANG, X., CHENG, K., SHI, F., JIANG, L., ZHANG, Y. & DOU, J. 2014. MiR-7, inhibited indirectly by lincRNA HOTAIR, directly inhibits SETDB1 and reverses the EMT of breast cancer stem cells by downregulating the STAT3 pathway. *Stem Cells,* 32**,** 2858-68.

ZHANG, J.-G., WANG, J.-J., ZHAO, F., LIU, Q., JIANG, K. & YANG, G.-H. 2010. MicroRNA-21 (miR-21) represses tumor suppressor PTEN and promotes growth and invasion in non-small cell lung cancer (NSCLC). *Clinica Chimica Acta,* 411**,** 846-852.

ZHANG, S., WU, T., PENG, X., LIU, J., LIU, F., WU, S., LIU, S., DONG, Y., XIE, S. & MA, S. 2017. Mesenchymal phenotype of circulating tumor cells is associated with distant metastasis in breast cancer patients. *Cancer management and research,* 9**,** 691-700.

ZHAO, Y., ALTENDORF-HOFMANN, A., POZIOS, I., CAMAJ, P., DÄBERITZ, T., WANG, X., NIESS, H., SEELIGER, H., POPP, F., BETZLER, C., SETTMACHER, U., JAUCH, K.-W., BRUNS, C. & KNÖSEL, T. 2017. Elevated interferon-induced protein with tetratricopeptide repeats 3 (IFIT3) is a poor prognostic marker in pancreatic ductal adenocarcinoma. *Journal of Cancer Research and Clinical Oncology,* 143**,** 1061-1068.

ZHUANG, Y. & MISKIMINS, W. K. 2011. Metformin Induces Both Caspase-Dependent and Poly(ADP-ribose) Polymerase-Dependent Cell Death in Breast Cancer Cells. *Molecular Cancer Research,* 9**,** 603-615.

ZI, F., ZI, H., LI, Y., HE, J., SHI, Q. & CAI, Z. 2018. Metformin and cancer: An existing drug for cancer prevention and therapy. *Oncology letters,* 15**,** 683-690.